INTRODUCTION TO

Physical Optics

BY

JOHN KELLOCK ROBERTSON, F.R.S.C.
Professor of Physics, Queen's University,
Kingston, Canada

THIRD EDITION—THIRD PRINTING

NEW YORK
D. VAN NOSTRAND COMPANY, INC.
250 FOURTH AVENUE

First Published, September 1929
Reprinted, March 1930, December 1931, February 1933, April 1934

Second Edition, August 1935
Reprinted, January 1937, February 1938, October 1939, September 1940

Third Edition, April 1941
Reprinted, February [illegible], February [illegible]

PRINTED IN THE UNITED STATES OF AMERICA
BY THE PLIMPTON PRESS · NORWOOD · MASS

DEDICATED

JOHN CUNNINGHAM McLENNAN

O.B.E., Ph.D., D.Sc., LL.D., F.R.S.

TEACHER AND INVESTIGATOR,

WHO FIRST SHOWED THE AUTHOR

THE BEAUTY OF THE STUDY OF

NATURAL PHILOSOPHY

PREFACE TO THIRD EDITION

SINCE the second edition of this book was printed, Prof. Rı Nielsen's committee on the Teaching of Geometrical Optics h: published its valuable report. Although I have no intention (altering the general nature of a text in which Physical Optics ar physical rather than geometrical ideas are stressed, in the thiı edition use has been made of some of the recommendations in th report. This has been all the more desirable, because, at the su; gestion of Prof. Bernhard Kurrelmeyer of Brooklyn College, f(the first time material on the microscope has been introduced.

The section on Polarization and Double Refraction has bee altered so that the uses, advantages and disadvantages of Polaroi are discussed in the logical development of this work. In this se(tion new demonstration experiments have also been introduced.

The revision has given the author an opportunity of amplifyin one or two sections in the chapters on Interference and Diffractior and of making a few minor corrections. A certain number of ne problems without answers have been added at the end of many (the chapters.

In making the revision I have taken into consideration sugge tions and criticisms which users of the text have been kind enoug to send me. For these and for words of appreciation, it is a plea ure to thank Prof. Bernhard Kurrelmeyer, Prof. D. L. Eator Prof. J. R. Pratt, Prof. W. W. Steffey, Miss A. W. Foster an Mr. B. J. Wender. If the names of any others are omitted fror this list, it is due to oversight, not to any lack of appreciation.

I should like also to thank the Bausch and Lomb Optical Con pany for literature kindly sent me.

J. K. R.

February, 1941.

PREFACE

In the field of physical optics, the number of books which, neither elementary nor advanced in character, provide a comprehensive introduction to the subject, is singularly limited. As the rapid advances of the last decade or two have accentuated the need for such books, the author feels that no apology is needed for placing this one before the public. In doing so he has had in mind the needs of two classes of students: (1) those who, at the outset of an intensive study of physics, are laying a thorough foundation for subsequent work in the theory of optics; (2) those specializing in other branches of science, for whom a general knowledge of modern views of light is desirable and, indeed, frequently indispensable. It is hoped, too, that the treatment is such that an appeal may be made to the general reader who desires to have some acquaintance with the fascinating problems of modern physics — problems many of which are most directly approached through the study of light.

In developing his subject the author has tried to steer a middle course between the Scylla of extreme conservatism and the Charybdis of radicalism. The almost perfect agreement between so many facts and the predictions of a wave theory, the constant necessity in light of speaking of wave-lengths, the manner in which frequencies are so intimately connected with quanta — all these things and many others demand a thorough discussion of wave motion and its light ramifications. A large part of the book, therefore, is devoted to such a study. But it would be conservatism indeed not to give an important place to quantum ideas. While (except for a brief reference at the end of Chapter I) no mention is made of the quantum theory until the book is well advanced, it will be found that Chapter XVII paves the way naturally for the introduction of this branch of the work. Moreover, in the writer's opinion, the historical development of the subject provides also the logical method for introducing the quantum theory, and the earlier part of the book is a necessary background for this purpose.

In presenting this branch of the work, the author has kept in mind the rapidity with which views have changed regarding

details of quantum theories, and has tried to avoid the danger leading the student to accept as final an outlook which is possib only a passing phase in the search for something more fundament

While the author lays no claim to any striking originality method, in his opinion the book is characterized by a continui which gives the subject a desirable unity, and by the stress laid a quantitative discussion of many questions. Thus the somewh extended discussion of wave motion in Chapter II, with its freque numerical examples, is justified on the ground that hazy ideas a best avoided by quantitative methods. Much use has also be made of graphical methods by means of which physical ideas a kept uppermost in the student's mind.

The preparation of the book has been greatly assisted by th clerical and other assistance provided by Dean A. L. Clark, whose suggestion it was written, and by the authorities of Queen University. It is the pleasant duty of the author to acknowled his indebtedness to them, as well as to all others who were of assi tance. His grateful thanks are due his wife for constant he throughout the preparation of the manuscript; to his form teacher, Dr. J. C. McLennan, of the University of Toronto, f photographs illustrating the Zeeman effect; to Dr. J. Stuart Foste of McGill University, Montreal, for the gift of a photograph of th Stark effect; to Dr. R. S. DeLury, of the Dominion Observator Ottawa, for photographs of solar and sunspot spectra, portions which are reproduced in the book; to Dr. F. H. Mohler, of the U. Bureau of Standards, for permission to reproduce the spectra magnesium shown in Plate II, Fig. 1; to the Editor of the Astr physical Journal for permission to reproduce Fig. 2; to the Sign Corps of the U. S. Army for permission to reproduce figures fro their radio manual; and to two former students, Mr. B. W. Sarger M.A., for taking the nitrogen band spectrum shown in Plate Fig. 3 and Mr. J. H. Findlay, M.Sc., for help in taking some the diffraction photographs. The numerous line diagrams we drawn by Mr. E. Harris, of Queen's University, and his nev failing care with this laborious task is greatly appreciated.

J. K. R

Queen's University,
Kingston, Canada
September, 1929.

CONTENTS

LIST OF PLATES

UNIVERSITY PHYSICS SERIES.

A Group of Text Books for Intermediate and Advanced Courses.

LEIGH PAGE, PH.D.
CONSULTING EDITOR.

Introduction to Physical Optics. *By* John K. Robertson, F.R.S.C., *Professor of Physics, Queen's University.*

Principles of Electricity. *By* Leigh Page, Ph.D., *Professor of Mathematical Physics, Yale University;* and Norman Ilsley Adams, Jr., Ph.D., *Assistant Professor of Physics, Yale University.*

Physical Mechanics. *By* Robert B. Lindsay, Ph.D., *Associate Professor of Theoretical Physics, Brown University.*

TEXTBOOKS ON PHYSICS

By

JOHN KELLOCK ROBERTSON, F.R.S.C.

INTRODUCTION TO PHYSICAL OPTICS
Third Edition.

Numerous diagrams and photographs clarify the presentation of this important subject.

Among the up-to-date topics treated at length are the optical methods for measuring the velocity of light, doublet and triplet series relations, the vector model of the atom, and the importance of isotopes in spectroscopy.

ATOMIC ARTILLERY

In this book the latest and most authentic information about the great discoveries of recent years in atomic physics, is fully presented by a distinguished scientist. He tells about electrons, protons, positrons, photons, neutrons and cosmic rays.

RADIOLOGY PHYSICS

An Introductory Course for Medical or Premedical Students and for all Radiologists.

PUBLISHED BY

D. VAN NOSTRAND COMPANY, INC.

NEW YORK

PLATE I

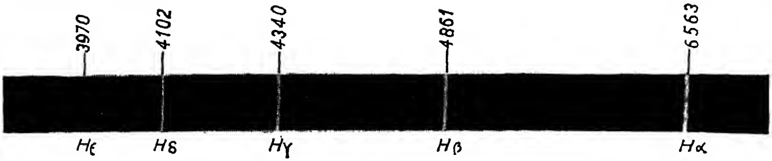


FIG. 1. Vacuum tube spectrum of hydrogen. (Glass Prism Spectrograph.)

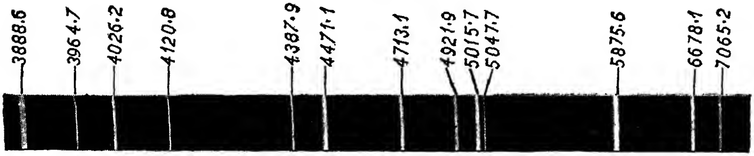


FIG. 2. Vacuum tube spectrum of helium. (Glass Prism Spectrograph.)

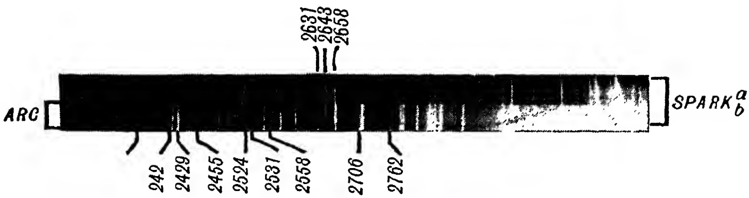


Fig. 3. A portion of the ultra-violet spectrum of tin, spark overlapping arc.

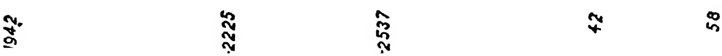


FIG. 4. Ultra-violet spectrum of mercury (Quartz Spectrograph.)

CHAPTER I

WHY WE STUDY WAVE-MOTION

1. Light may be provisionally defined as the physical agency by means of which the eye receives the sensation of sight. In this book it is proposed to discuss the nature of this physical agency, and, if possible, to give some answer to the question, What is light? Such a discussion constitutes what we call Physical Optics, although in the course of our study it will be seen that our definition of light is too restricted to include all the work of that subject.

To give any kind of answer to the question we have raised, it is necessary first of all to consider facts. In one sense, therefore, it may seem absurd to attempt any discussion of the nature of light until we have completed our study of its properties. On the other hand, there are certain outstanding simple facts so important that, in the light of them, we may formulate tentative theories to be subsequently tested as additional information is acquired. First of all, then, the student is asked to consider the following facts with most of which he has probably long been familiar.

(*a*) *Matter is not necessary* for the propagation of light. While light passes readily through certain kinds of matter which we call transparent, its passage is unobstructed in the vast interstellar spaces, or through a glass vessel from which all the air has been exhausted. Contrast this fact with the case of sound which is not propagated in a vacuum.

(*b*) Light is *regularly reflected* at a smooth surface such as a mirror, subject to the simple law that the angle of reflection is equal to the angle of incidence. This fact was known as long ago as the time of the Platonic school, which flourished four centuries before the Christian era.

(*c*) Light is *refracted* at a surface separating two transparent media of unequal density, such as air and water. In the case of a plane surface, the sine of the angle of incidence bears a constant

ratio to the sine of the angle of refraction. (Snell's Law.) While the fact of refraction was known, and observations were made as far back as the time of Ptolemy (Greece, 2nd Cent. A.D.), the sine law was not discovered until early in the seventeenth century, probably independently by Snell (Holland, 1591–1626) and by the famous philosopher Descartes (France, 1596–1650).

(*d*) Light may be *simultaneously* reflected and refracted. For example, if a beam of light such as 1, Fig. 1, be allowed to strike, at a suitable angle, the under surface *AB* separating air from water, both a refracted beam 2 and a reflected beam 3 are observed.

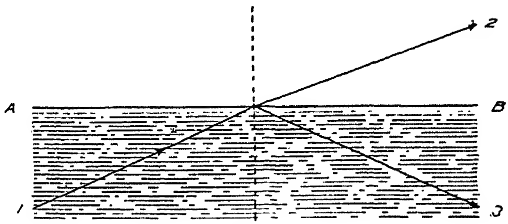


FIG. 1.

(*e*) For all practical purposes, light travels in straight lines — the fact of rectilinear propagation. This, too, was known to the Greek philosophers. Of its approximate truth, the sharpness of the shadows of obstacles cast by a small source of light, as well as the practical use to which it is put in placing objects in a straight line by "sighting," give ample evidence.

It may here be noted that under special circumstances which we shall later consider, it may be shown that light does not always travel in straight lines, a fact to which attention was first drawn by Grimaldi (Italy, 1618–1663).

(*f*) Light travels in free space at a very definite velocity, which, within 0.06 per cent, is equal to 3×10^{10} cm. per sec. In section 4A at the end of this chapter a detailed description will be given of some of the methods used for determining this important constant. Here we merely wish to call attention to this very important fact.

(*g*) The velocity of light in a medium such as water is *less* than the velocity in air. This important fact was first experimentally shown by Foucault (France, 1819–1868) about 1862 by a method to be described in section 4A. Some years later, in 1883, Michelson (United States, 1852–1931) showed that the ratio of the velocities for these two media was 1.33.

(*h*) Energy is carried by a beam of light. The student, who scarcely needs to be reminded of this fact, will at once recall the rise in temperature which results from the absorption of the sun's rays, as well as other manifestations of energy due to the same

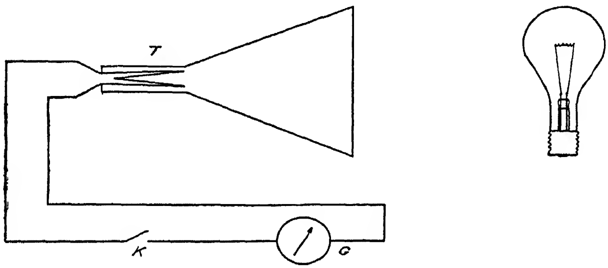


FIG. 2.

cause. The fact may be demonstrated otherwise by a simple experiment such as is represented in Fig. 2, where *T* is a thermopile, *G* a galvanometer, *K* a knife-key, and *L* an incandescent lamp so placed that rays from it may fall on one juncture of the thermopile. On closing the key *K*, the galvanometer indicates an electric current which continues as long as the lamp is lighted. Energy, therefore, passes from the lamp to the thermopile.

2. Two Theories of Light. — These facts may be supplemented by many others. Those we have given, however, are outstanding, and for that reason form a good basis for a preliminary discussion concerning the nature of light. If we take the last statement first, it will readily be evident that only two general theories are possible, because energy can be transmitted from one place to another by only one of two general means. A window pane may be broken

from a distance either by being struck by some moving object, or by the concussion resulting from a disturbance which has travelled through the atmosphere from the region of an explosion. We hear a distant source of sound because of the impact on the ear of waves which have travelled through the air from the source to the ear. We detect an odour from an object at some distance away by the motion of particles which have emanated from the object. In general, energy can be propagated either by moving matter or by a wave disturbance travelling through a medium which does not itself move as a whole. In light, therefore, two general theories have been formulated. According to the one, *the wave theory*, a disturbance travels from a light source through a surrounding medium; differences in colour corresponding to differences in wave-length; according to the other, *the emission or corpuscular theory*, light consists of a flight of invisible rapidly moving particles whose size varies with the colour, projected from the luminous source. One important difference between the two theories may here be noted. On the wave theory, at all points on a surface through which an ordinary beam of light is passing, energy is *uniformly and continuously* distributed; whereas, on the other, the energy distribution is *discontinuous*, being concentrated at points. In the seventeenth and eighteenth centuries, when the pioneers of modern science were laying the foundations of the structure which to-day is so elaborate, each of these theories received strong support. We find a wave theory upheld by such men as Hooke (England, 1635–1703), Huygens (Holland, 1629–1695), Descartes (France, 1596–1650), and Euler (Switzerland, 1707–1783); while the corpuscular theory was defended by such intellectual giants as Newton (England, 1642–1727) and Laplace (France, 1749–1827). Which, then, is the better?

3. Corpuscular *versus* Wave Theory. — A theory is satisfactory only in so far as it provides an explanation of facts. A single fact for which a satisfactory explanation cannot be given, or which goes contrary to a deduction which may reasonably be made from the theory, weakens it. If many such facts come to light, the

theory has to be discarded, or at least radically altered. In testing two rival theories, therefore, we must examine carefully as many facts as we can to see which provides the more satisfactory explanation. We begin naturally with the list of facts which has been enumerated. In the light of these, is it possible to give a decided preference to one theory rather than the other? Let us consider the facts one by one.

(*a*) As it is impossible for most minds to think of waves without a medium to carry the wave-motion, and as matter is not necessary for light propagation, a hypothetical luminiferous ether has to be postulated. Although no such postulate is necessary on the corpuscular theory, this was never considered an objection to the wave theory. As will be seen presently, the assumption of an ether in which vibrations may be set up played an important part in the corpuscular explanation given by Newton for more than one fact. To quote from Newton's writings, "It is to be supposed that there is an ethereal medium, much of the same constitution with air, but far rarer, subtiler, and more strongly elastic."

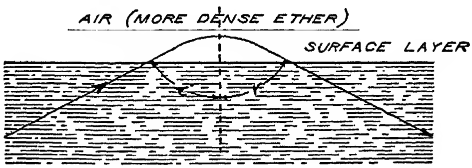


FIG. 3.

(*b*) and (*c*) As the explanation of both reflection and refraction on the wave theory will be given in some detail in Chapter III, we shall here consider only that given by Newton on a corpuscular theory. In this the supposition of an ether of variable density, *more* dense in free space than in solids, and of decreasing density at the surface of a body, plays an important part. When a corpuscle strikes a surface separating two media, therefore, it encounters a

region of variable density — of "graduated rarity," as Newton put it. The explanation involves the further assumption that the speed of the corpuscle is accelerated in passing from more dense ether (*e.g.* air) to less dense (*e.g.* glass), and retarded when travelling in the opposite direction. In the case of *total* reflection, the retardation is sufficient to gradually turn ("incurve") the particle in its passage through the surface layer, somewhat as illustrated in Fig. 3, so that it emerges in a direction making the angle of reflection equal to the angle of incidence. In refraction, on the other hand, "if the differing densities of the mediums be not so

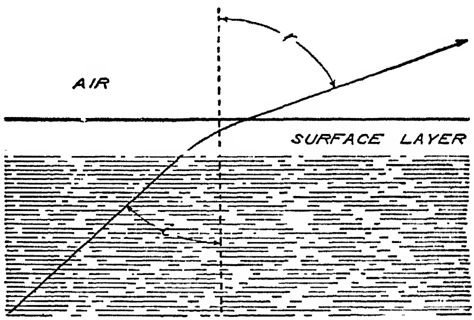


FIG. 4.

great, nor the incidence of the ray so oblique as to make it parallel to that surface (superficies) before it gets through, then it goes through and is refracted" — somewhat as shown in Fig. 4. *It is important to note that this explanation of refraction involves the assumption that a ray of light is retarded, that is, travels at a less rate in air than in a medium like water.*

Ordinary reflection, such as at a surface separating air from glass, is accounted for by assuming (apparently because of the peculiar surface effects which we now study under surface tension) that "ether in the confine of two mediums is less pliant and yielding than in other places," in consequence of which, "a ray finds more than ordinary difficulty to get through," and has its direction reversed just as when a corpuscle encounters denser ether.

To the twentieth-century student such explanations may seem very artificial, but if he considers how readily a beam of invisible cathode rays may nowadays be deflected from a straight line path by an electric or a magnetic field, he will recognize that something is to be said for the corpuscular view-point.

(*d*) The exponents of the corpuscular theory were by no means unaware that reflection and refraction may take place simultaneously at the same surface. On the wave theory this is to be expected (see Chapter III). In the light of the corpuscular explanation which has just been given, how can some particles be sent back from the surface while others are transmitted? To that question Newton gave an ingenious answer by means of his theory of "fits of easy reflection and easy transmission." According to this, "the rays when they impinge on the rigid resisting ethereal superficies . . . cause vibrations in it, as stones thrown in water do in its surface; and these vibrations . . . *alternately contract and dilate the ether in that physical superficies*." In consequence of this, the surface layer of ether, when contracted, is sufficiently dense to cause reflection of the ray, whereas, when dilated, it is sufficiently rarefied to allow transmission, that is, to give rise to refraction.

It will be seen, then, that the corpuscular theory involved the idea of waves in an ether, even if rays themselves were considered corpuscular. It would seem, therefore, that preference should be given the simpler wave theory, provided it gives at least an equally satisfactory explanation of other facts.

(*e*) There was one important fact, however, the explanation of which on the wave theory was long delayed, and that was the fact of approximate rectilinear propagation. This was the big stumbling-block to those early scientists who were unable to accept the wave theory. How can waves travel in straight lines, they said? A whistle can easily be heard around the corner of an obstacle; a candle certainly cannot be seen from behind it. The one is undoubtedly a wave phenomenon, how can the other be? And so we have the objection given in the following words of Newton: "If it (light) consisted in pression or motion . . . it would bend

into the shadow. For pression or motion cannot be propagated in a fluid in right lines, beyond an obstacle which stops part of the motion, but will bend and spread every way into the quiescent medium which lies beyond the obstacle . . . ."

On the other hand the emission theory provided a ready explanation of rectilinear propagation. This fact, combined probably with the weight given to the opinion of such an intellectual giant as Sir Isaac Newton, was largely responsible for the preference given the corpuscular theory until the early part of the nineteenth century. That there was a slight deviation from the straight line path when, for example, a beam emerging from a small hole, passed

FIG. 5.

by the edge of an obstacle, was known in Newton's time. But this bending was slight and at that time a satisfactory explanation of the phenomenon on the wave theory was not forthcoming, although both Hooke and Grimaldi were on the verge of discovering the principle which more than a century later was to clarify the whole question.

The satisfactory explanation of rectilinear propagation on the wave theory is largely the result of the work of Thos. Young (England, 1773–1829), and of Fresnel (France, 1788–1827), both of whom demonstrated the phenomenon of interference of light, and showed its implications. In Chapters VIII and IX this subject will be discussed in detail. In view, however, of the question now under discussion — the relative merits of the corpuscular and the wave theories — it is necessary at this stage to consider briefly its significance. This can perhaps best be done by reference to a simple experiment which the student should try for himself.

If two narrow slits, S_1 and S_2, (Fig. 5), about 1 mm. apart, are cut in a piece of dark paper, or scratched on a bit of old photographic

negative, and an observer views through them a narrow source of light S, such as a straight filament lamp, placed two or three metres away, a series of alternately dark and light narrow bands is seen in the centre of a tolerably wide band of light. Should a red glass be interposed between the source S and the slits, a series of red and black narrow bands is observed. Some idea of their appearance will be obtained from Fig. 1, Plate IV, a reproduction of an actual photograph taken with an ordinary lens replacing the eye lens, and a sensitive plate, the retina. Now what is the significance of these alternately light and dark bands? A wave theory of light, combined with the principle of interference, gives the answer.

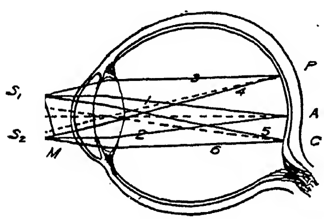


FIG. 6.

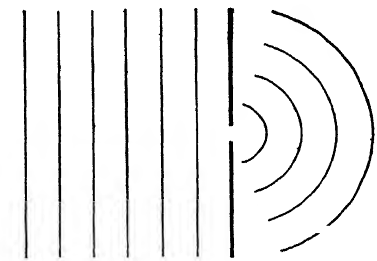


FIG. 7.

If light is a wave phenomenon, then it is reasonable to assume (what indeed can experimentally be shown) that rays spread out to some extent from each slit S_1, S_2 (Fig. 6) just as water waves spread out from a narrow opening, somewhat as represented in Fig. 7. The eye lens then collects these rays, and, since the retina is approximately at its focal distance from the lens (when a distant source is viewed), parallel pairs of rays are brought together at different points on the retina. For example, if the student recalls the simple lens law, that a ray goes through the centre unchanged in direction, he will see that rays 1 and 2 will meet at A, rays 3 and 4 at P, rays 5 and 6 at C and so on. Now, if rays are corpuscles, wherever they strike a surface, whether there are two bundles or only one, there must be light. But what if rays are waves?

Consider the point P. The path S_2 to P is longer than the path

S_1 to P by the distance S_2M, where S_1M is perpendicular to S_2M. On the wave theory we have to do with what are analogous to crests and troughs in water waves. If, therefore, S_2M is half a wave-length (distance crest to trough), or any odd number of half wave-lengths, it is evident that at P crests from S_1 always meet troughs from S_2, and vice versa, whereas if S_2M is one, two, three or any whole number of wave-lengths, crests always meet crests, troughs, troughs. Now, according to the principle of interference, "when two undulations from different origins coincide . . . their joint effect is a combination of the motions belonging to each." If we accept this, the joint effect of a crest and a trough is zero, or, in other words, there is, in such a case, no disturbance and no light, whereas at places where two crests or two troughs meet, the joint effect produces a maximum disturbance, that is, bright light. The explanation, therefore, of the alternate dark and light bands is simply that, because of the varying difference in the paths from the two slits to successive points on the retina, we have alternately reinforcement and annulment of the two sets of waves. In Chapter VIII this experiment will be discussed quantitatively, and it will then be seen that with the aid of simple linear measurements it provides us with a means of evaluating the actual lengths of light waves.

The principle of interference when first published about 1801 by Young did not receive general acceptance. Indeed, in some quarters it was almost ridiculed. When a few years later, however, Fresnel showed how by means of it a ready explanation could be given of approximate rectilinear propagation, as well as of the observed slight bending to which reference has been made, the wave theory gradually came into great favour. It was seen to provide a much simpler explanation of optical facts than the corpuscular.

Acceptance of the wave theory became almost universal when Foucault and Michelson showed experimentally that the speed of light is less in water than in air. This was a fact which provided a crucial test of the relative merits of the two theories, for the explanation of refraction on the corpuscular theory involved, as we

have already noted, the assumption of a greater velocity in water than in air, whereas the wave theory postulated a less velocity. In the light of this experimental fact, the corpuscular theory could not be accepted, at any rate without radical revision, and the wave theory took complete possession of the field. By means of it, throughout the last half of the nineteenth century tremendous advances were made in the interpretation of light phenomena. The form of the theory was not always the same, the conception of vibrations in an elastic solid ether, for example, giving way to Maxwell's electromagnetic theory (Chapter XVI), but it was always a wave theory. Indeed, so perfect was the agreement between fact and theory that in 1889 we find Hertz (Germany, 1857–1894) making the statement, "The wave theory of light is from the point of view of human beings, a certainty."

4. That, however, was a rash statement for any scientist to make, and this chapter would not be complete without a brief reference to some work of the twentieth century which should warn us against the too hasty acceptance of any theory. For a generation the attention of physicists has been directed to problems connected with the emission and the absorption of light. This work has shown that there are many facts for which an adequate explanation cannot be given in terms of a simple wave theory. One of the most important of these facts, strangely enough, was discovered by Hertz in the very year in which the above statement was made — the fact that light when incident on a metal plate may cause an emission of electrons. When the laws governing this process were later discovered, it was found that a *quantum theory* which postulates the emission of light *discontinuously* in isolated bundles or *quanta* provided an explanation, whereas the continuous wave-front of the wave theory failed to do so. A *quantum* is a definite unit of radiant energy equal to the product of h, a certain universal constant about which more will be stated later, and, ν, the frequency of the radiation.

There are certain phenomena, therefore, which can be interpreted only by thinking of light as having a corpuscular nature,

or a kind of atomicity, the fundamental unit possessing a quantum of energy. This unit is frequently called a *photon*. In many ways this is suggestive of the old corpuscular theory but it is by no means a return to it. Photons are not material particles, as Newton's light corpuscles were considered to be. Moreover, the mere fact that the frequency of radiation is involved in the numerical magnitude of the energy of a photon shows that there is an intimate connection between quantum and wave theories. We must not forget, too, that other phenomena such as interference and diffraction find a complete and satisfactory explanation by an ordinary wave theory.

The modern position, then, is that light has a dual nature. Sometimes we must think in terms of photons, sometimes in terms of waves. To this whole question we shall return in later chapters, but, at this stage, it should be evident that no intelligent study of light is possible without a clear understanding of wave-motion in general. In the next chapter this is the subject of discussion.

4A. Methods of Determining the Velocity of Light.

(*a*) **Galileo's Attempt.** — The direct method of determining any velocity, either of a particle or of a disturbance, is to observe the time taken to pass over a measured distance. It is not surprising, therefore, to find that Galileo (Italy, 1564–1642) tried in a simple, direct, if crude, way to observe the time taken for a beam of light to travel a distance of a few miles. Two observers, A and B, who had previously practised uncovering the light from lanterns which they held at close range, until they acquired "such skill .... that the instant one sees the light of his companion he will uncover his own," were subsequently stationed a few miles apart at night. A was then to note the time interval which elapsed between the moment he uncovered his lantern and the instant he saw B's light. This interval obviously is the time taken for light to travel from A to B and back again to A. The experiment failed because the time taken to cover a distance of a few miles by a disturbance travelling at a speed of 3×10^{10} cm. per sec., or 186,000 miles per second, is far too short to be measured in such a crude way. The

principle of the experiment, however, is exactly the same as that utilized in the refined methods described below.

(*b*) **Römer's Astronomical Method.** — To make measurements of such a high velocity, either the measured distance must be so great that time values can be easily observed or methods of measuring extremely small intervals of time must be devised. Conditions for the first alternative are supplied by astronomical distances, and the earliest determination of the velocity of light was made

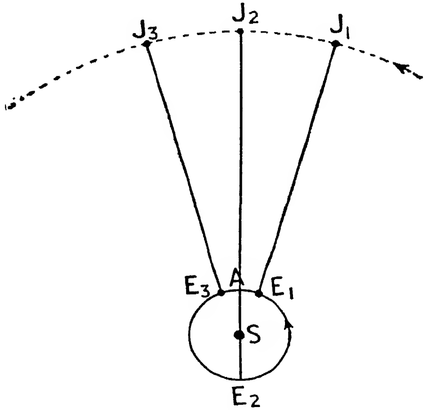


Fig. 8.

in this way by Römer (Denmark, 1644–1710), when working in Paris, in 1676. To understand his method, it is necessary to recall that the earth and Jupiter are both planets revolving about the sun, the earth with its yearly period, Jupiter with a period just under 12 years. It follows, then, that in a little over half a year (0.545 of a year, to be more exact) the earth will pass from a position of conjunction, where it is nearest Jupiter, as at E_1, Fig. 8, to a position of opposition, as at E_2, where the distance between these planets is a maximum. It is also necessary to know that Jupiter has several moons or satellites which, to an observer on the earth, are eclipsed when they get behind the planet. Confining

our attention to a single satellite (for example, to one which makes a complete revolution about Jupiter in less than two days), we see that it is a simple matter, by observing the times at which this satellite is eclipsed, to determine its period of revolution. It is then possible to predict the time of future eclipses. In Römer's time this was done and it was noticed that there were unexpected disagreements between the observed and the predicted times of eclipses. In 1676 Römer reported to the French Academy of Science that a certain eclipse would not occur until ten minutes after the calculated time, and, at the same time, stated that the cause of the disagreement was the fact that light travels with a finite velocity. The calculated results were out, he said, because this factor had been neglected in the calculations. By reference to Fig. 8 it will be seen that, when the earth is in position E_2, light has to travel from Jupiter a distance J_2E_2 which is greater than J_1E_1, the distance when the earth is in position E_1, by an amount equal to the diameter of the earth's orbit about the sun, that is, by some 196 million miles. If we take this fact into consideration, the velocity of light can be calculated in the following way.

Let t = the actual time of one revolution of a satellite,
n = the number of revolutions made by the satellite during the time that the earth moves from the position E_1 to E_2, that is, during the interval conjunction to opposition.

Then T_1, the *observed* time for these n revolutions, is given by

$$T_1 = nt + \frac{D}{c},$$

where D = the diameter of the earth's orbit,
c = the velocity of light in free space.

Similarly, if T_2 is the observed time of the n revolutions which take place during the interval from opposition to conjunction, that is, when the earth moves from E_2 to E_3,

$$T_2 = nt - \frac{D}{c}.$$

Hence $$T_1 - T_2 = \frac{2D}{c}.$$

Now $T_1 - T_2$ is an observed quantity. Using the value obtained from the observations of his time and the astronomical estimate of the size of the earth's orbit, Römer found a value for c, the velocity of light, not very different from the standard value of to-day. (Now that the value of this quantity has been obtained to a high degree of accuracy, the modern value of $T_1 - T_2$, which is 16 min. 26 sec., may be used to evaluate the mean diameter of the earth's orbit.)

(*c*) **Bradley's Aberration Method.** — The essential ideas underlying this second astronomical method, first noted in 1728 by Bradley (England, 1693–1762), are discussed in another connection in the last chapter of this book.

(*d*) **Fizeau's Toothed-Wheel Method.** — It was not until 1849 that the velocity of light was measured by what we might call a

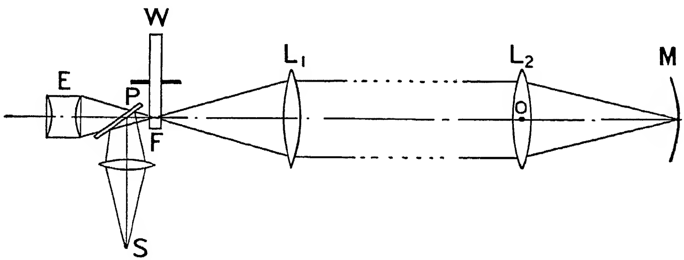


FIG. 8A.

laboratory method, due to Fizeau (France, 1819–1896). In this, as in all similar methods, astronomical distances are replaced by a length of a few miles and a device is used for measuring an extremely short time interval. The essentials of Fizeau's arrangement are shown in Fig. 8A. Light from a small source S by means of a lens and reflection from one face of a transparent plate P is focussed to form a small spot at F from where it diverges, falling on the lens L_1 which makes the beam parallel. After traversing a distance, which in Fizeau's original experiment was 8633 metres,

the parallel beam passes through the lens L_2 which focusses it on the face of a concave mirror M, so placed that its centre of curvature coincides with O, the centre of lens L_2. The beam of light is then reflected back along the incident path. On striking the transparent plate P part of the light is transmitted and so the spot F can ordinarily be observed in an eyepiece E. W represents a toothed wheel, with 720 projecting teeth, separated by open spaces of equal width.

Suppose, now, with the wheel in rapid rotation, that at a certain instant the spot F is exactly between two teeth, somewhat as shown in Fig. 8B(*a*). As the light which leaves the spot at this instant travels to M, the wheel is turning and, for a particular

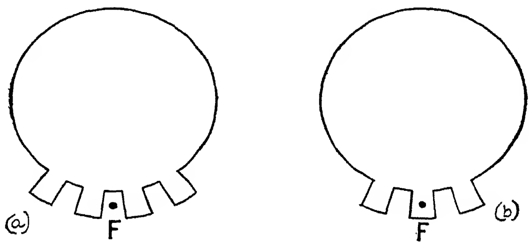


FIG. 8B.

speed of rotation, the reflected beam, on arriving back at F, will strike the centre of a tooth, as in Fig. 8B(*b*). If the wheel is maintained at this speed, the light which gets through an opening will always arrive back at the wheel in time to strike a tooth, and so for this speed of the wheel an eclipse is noted by an observer looking through E, Fig. 8A. At twice this speed of rotation, the light on its return to F strikes the centre of the next opening and, in this case, is seen as a bright spot. By still further speeding up the wheel, alternate eclipses and bright spots can be seen. From observations of the value of these critical speeds, a simple calculation gives the speed of light. For example, in one of Fizeau's experiments, the first eclipse occurred at a speed of 12.6 revolutions per second. Therefore, the time taken to travel twice 8633 metres is equal to $\frac{1}{2} \cdot \frac{1}{720} \cdot \frac{1}{12.6}$ seconds, from which the velocity of light

is at once found. As in Table A, page 22, values of this constant are given for many of the exact determinations in recent years, definite values are not given for each of these earlier methods.

In 1874–1878 Fizeau's method was repeated by Cornu (France, 1841–1902), with the distance increased to 23,000 metres, with an improved method of measuring time, and with observations made of the change in speed when the brightness of F passed from a certain intensity (observed by comparison with a fixed source) through an eclipse back to this same intensity.

(*e*) **Rotating Mirror. Displaced Image Method.** — A method suggested in 1838 by Arago (France, 1786–1853) was planned

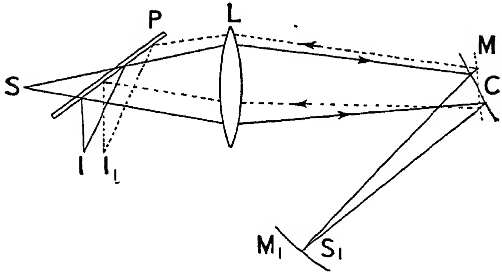


FIG. 8C.

jointly by Fizeau and Foucault and subsequently carried out independently by Foucault, who reported his results in 1862. The principle of the method can be understood by reference to Fig. 8C.

Light emerging from a narrow source S passes through a transparent plate P, falls on a lens L, is converged to a plane mirror M, capable of rotation, which reflects the beam to the concave mirror M_1. Here an image of the source S_1 is formed. As the centre of curvature of this concave mirror is at C (a point on the axis of rotation of the plane mirror M), the light normally retraces its path back to the source S. By means of the transparent plate P part of the returning beam is reflected to form an image at I. *Even if the position of the plane mirror M is altered* so that the incident beam strikes it at a different angle, the image will still be formed at I because, although the position of S_1 varies

with the inclination of M, the reflected beam between M_1 and M always follows the incident path, and so strikes M at the same angle as it left M. Suppose, however, that the mirror M is rotated so quickly that, during the time it takes for light to go from this mirror to M_1 and back again, M has moved to the position represented by the dotted line. In that case, the return beam strikes the face of the plane mirror at an angle which is not the same as the angle with which the light left this mirror. The reflected beam, therefore, no longer retraces its path back to I but follows a slightly different path, giving rise to a displaced image at I_1. The magnitude of the displacement II_1 evidently depends on the angle through which M is rotated while light travels twice the distance M to M_1. By observing the value of II_1, therefore, this angle can be found, and hence, from the speed of rotation of the mirror, the corresponding small interval of time. From this time value and twice the distance MM_1, the velocity of light is at once obtained.

In Foucault's arrangement MM_1 was an effective length of 20 metres. (The actual length was only 4 metres, but the beam was reflected back and forth five times between fixed mirrors.) The displacement II_1 was only 0.7 mm. but the images were sharply defined and their positions could be located with precision.

It was by this method that Foucault proved more or less qualitatively the very important result that the velocity of light in water is less than in air. By the same method Michelson proved that the ratio of these two velocities was 1.33.

In the period 1879–1882 the rotating mirror method was repeated with great care by Michelson and by Newcomb (United States, 1835–1909). As it is our aim to give principles rather than refined experimental details, we shall simply point out that Michelson, using a path length of 600 metres, obtained an image displacement of 133 mm. and that Newcomb, working with a distance of 1200 ft., used a four-sided (cubical) mirror which increased the brightness of the image.

(*f*) **Rotating Mirror. Null Method.** — In the rotating mirror method the angle through which the mirror rotates cannot be

measured to the same degree of accuracy as the speed of rotation and the length of the optical path. This difficulty can be overcome by making use of a suggestion made by Newcomb that a many-sided mirror be used and an optical path of such a length that the return beam strikes one face of this mirror *at the same angle* as it left the adjoining face, on its outward journey. In that case, there is no displaced image. To avoid practical difficulties, the optical arrangement may be such that the return beam strikes a face when it is *parallel* to the incident face. Thus, in Fig. 8D

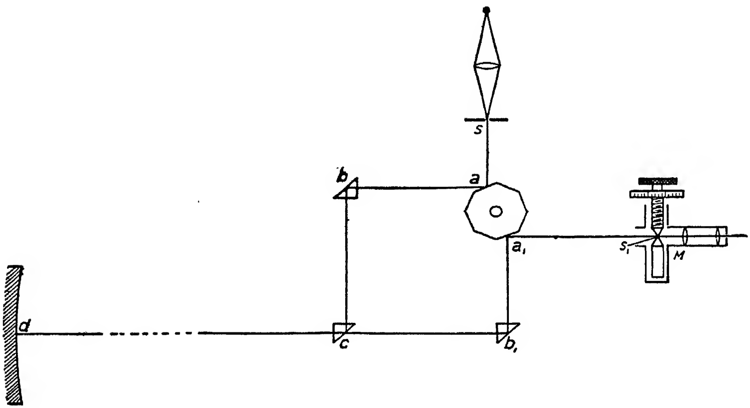


Fig. 8D.

(to be explained more fully presently), if the octagonal mirror there represented is at rest, the incident beam strikes one face at a and the return beam at a_1 on a parallel face. In applying this method with an octagonal mirror the mirror must be rotated until it is going at such a speed that it turns through 45° during the time that a beam of light has travelled from a to the outward reflector and back again to a_1. For that critical speed the beam is reflected from the last face along exactly the same path as when the mirror is at rest.

This is the principle used in an extremely painstaking investigation carried on under the direction of Michelson, during the years following 1924. One of the arrangements used will be clear from

Fig. 8D and from the following description, both of which are taken from a paper by Michelson in the *Astrophysical Journal*, 1924. "The light source is a Sperry arc which is focussed on the slit s. Falling on the face a of the octagon, the light is reflected to a right-angle prism b, to another at c, whence it proceeds to the concave mirror d, of 24-inch aperture and 30-foot focus. This reflects the pencil as a parallel beam to the distant mirror (also a 24-inch, 30-foot concave), proceeding thence to a small concave reflector at its focus. An image of the slit is formed at the face of this small reflector which necessitates the return of the light to the concave mirror at d, whence it passes *over* the prism c to b_1, whence it is reflected to the face a_1 of the octagon, forming an image at s_1, where it is observed by the micrometer eyepiece M."

In this experiment the light travelled from the observing station on Mt. Wilson to Mt. San Antonio, California, a distance of 22 miles, a length which was measured by the U. S. Coast and Geodetic Survey to an accuracy of nearly one part in seven million. The greatest care was also taken to measure the time with the utmost accuracy. The velocity of the revolving mirror was measured "by stroboscopic comparison with an electric fork" whose rate was obtained "by comparison with a free seconds pendulum, which last is compared with an invar gravity pendulum furnished and rated by the Coast and Geodetic Survey." By observing the slight difference in the positions of the images seen in the eyepiece when the mirror was revolved in opposite directions, a small correction was made for the inability to rotate at such a speed that the second face was *exactly* parallel to the incident face. Some idea of the amazing accuracy attained will be had from the following average results obtained, using not only an octagonal mirror, but others, both of glass and of steel, with 12 and 16 faces.

Glass 8 sides	299, 797	km. per sec. *in vacuo*.
Steel 8 "	299, 795	"
Glass 12 "	299, 796	"
Steel 12 "	299, 796	"
Glass 16 "	299, 796	"

(*g*) **The Tunnel Experiment.** — Still more recently, in the years following 1930–1931, the same general method has been employed by Michelson in collaboration with G. F. Pease of the Mount Wilson Observatory and F. Pearson of the University of Chicago. Since the death of Professor Michelson in 1931 this work has been carried on by Pease and Pearson. In this latest experiment an attempt was made to increase the accuracy by using a short base line which could be measured without triangulation. This was done by constructing a tunnel in the form of a galvanized pipe tube, one mile long, three feet in diameter, from which the air was exhausted by powerful vacuum pumps, to a pressure as low as 0.5 mm. of mercury. By working at such a low pressure, errors arising from the dependence of the velocity on the index of refraction of atmospheric air were eliminated, and, moreover, the sharpness of the optical images utilized was greatly increased. The base line was measured by the U. S. Coast and Geodetic Survey with a probable error of $\pm$ 0.47 mm., that is, to an accuracy of about one part in 3,400,000. With the aid of a 32-sided rotating mirror and an optical arrangement similar in principle to the one just described, all housed in evacuated chambers and operated by remote controls, the time taken for a beam of light to travel back and forth sometimes eight, sometimes ten times was measured with even greater care than in the 1925–1926 investigation. In the final report of this investigation the result given is

299,774 km. per sec. *in vacuo.*

This value is the simple mean of some 2885 determinations, but it is worth while noting that the averages of smaller numbers of determinations taken during different time intervals gave values differing from this final mean by more than experimental error. For example, the mean of one set of 55 determinations was 299,780 km. per sec., while the mean of another set of 47 (both sets being taken at different intervals in the period March to August, 1932) gave 299,771 km. per sec. The conclusion of Pease and Pearson that "attempts to explain these variations in velocity as a result of instrumental effects have not thus far been success-

ful" should be compared with the conclusion of Birge given at the end of the next section.

4B. Summary of Results of Velocity Experiments. — An examination of the values of the velocity obtained by the important methods of the last half century suggests certain periodic variations which are too great to be ascribed to experimental errors. This will be evident by a glance at Table A, below, which was published in a recent number of *Nature* (Nov. 17, 1934) by Professor Birge (United States). A careful analysis of these results suggests a periodic change, with a forty-year cycle. Moreover, according to Birge, "Pease and Pearson have found evidence of two shorter periods, one of 14¾ days, the other of one year." Although the experimental evidence thus points to *slight* variations in the value of c, it does not follow that these changes necessarily correspond to real variations in the value of c. In this connection it is particularly significant that no such fluctuations occur in the values obtained by the indirect methods which may be used to measure this quantity (see section 146). At the present time most physicists will agree with Birge's conclusion that the fluctuations in the values obtained by the direct methods, although not yet satisfactorily explained, are "instrumental rather than real."

Table A

Date	Investigator	Observed velocity
1874	Cornu-Helmert	299,990 ± 200
1879	Michelson	299,910 ± 50
1883	Newcomb	299,860 ± 30
1883	Michelson	299,853 ± 60
1902	Perrotin	299,901 ± 84
1926	Michelson	299,796 ± 4
1928	Mittelstaedt	299,778 ± 10
1932	Pease and Pearson	299,774

CHAPTER II

STUDY OF WAVE-MOTION

5. We shall begin our study of wave-motion by considering what may be learned from ordinary water waves with which every student is familiar. An observer sitting beside the shore of a lake and watching waves on its surface cannot fail to be struck with two things: (1) a floating object, such as a block of wood, executes a to and fro motion without altering appreciably its distance from the shore; (2) at any given instant, crests (places at each one of which the surface is displaced above its normal position a maximum amount) and troughs (places were the displacement is a maximum below the normal level) are repeated at regular intervals. If the observer wished to express these facts in more mathematical language, he might proceed, first of all, by taking an imaginary line along which the waves are travelling. He could then define the position of a particle at any instant by giving, (1) its distance x from some point of reference on his imaginary line (such as a stake in the water); and (2) the amount y the particle is displaced from its normal position. With changing time, y would alternate between a maximum plus value (at a crest) and a maximum minus value (at a trough). Our observer would then be able to re-state his observations as follows:

(1) When x is constant, y is a periodic function of the time, or

$$y = \text{periodic} f\,(t),\ x \text{ constant, and}$$

(2) when t is constant, y is a periodic function of x, or

$$y = \text{periodic} f\,(x),\ t \text{ constant.}$$

These two statements, which presently we shall combine into one, give the fundamental characteristics of a train of waves. They tell us that the same to and fro motion is handed on from particle to particle in the medium along which the disturbance is travelling.

At any instant, two particles whose times of vibration differ by an integral multiple of the periodic time are said to be in the same phase. Two consecutive particles in the same phase are separated by a *wave-length*.

6. The Cause of a Train of Waves. — Particles cannot be set vibrating in this manner without a cause. We must somewhere have a vibrating source. If, for example, we wish to send waves along a cord under tension, we first of all pluck aside a portion at one end. On releasing the cord the *restoring force* due to the component of the tension at right angles to the cord quickly brings the displaced portion back to its normal position. Because of its momentum, however, it does not remain there but overshoots the mark until a displacement is caused in the opposite direction. Once more the restoring force brings the displaced portion back, and so a to and fro motion is set up. A vibrating source has then resulted from the initial displacement. Generally speaking, this will be the case whenever a particle of finite mass, when displaced from its normal position, brings into play a sufficiently large restoring force.

The student will recall other examples. When a stone is thrown into a pond, a certain amount of water is displaced from its normal position, the restoring force of gravity is brought into play, and we have a vibrating source from which ordinary water waves spread out. If the displacement is very slight, the restoring force is largely the result of surface tension and we have the tiny waves called ripples. To have sound waves in air, by some means (*e.g.* a vibrating tuning fork) the air particles must be displaced from their normal position, thus bringing into play the restoring force of the elasticity of the air.

In this connection one further point should be noted. The restoring force which comes into play on displacement of a portion of a medium is the result of some sort of elastic connection between contiguous particles of the medium. For that reason a single particle cannot be displaced without some displacement of its immediate neighbor. It follows, therefore, since the displacement of any particle calls into play restoring forces, that the motion

of a vibrating source is going to be passed on from particle to particle, thus constituting a wave-motion.

7. A simple quantitative example may not here be out of place. *A source particle executing a periodic motion defined by $y = 6 \sin \pi t$ sends out waves which travel through a homogeneous medium at the rate of 6 cm. per second. Find the displacement of a second particle 80 cm. from the source one minute after the source began to vibrate.* (*Assume the medium of one dimension like a stretched cord.*)

Solution: Since the wave-disturbance travels at the rate of 6 cm. per second, the second particle will not begin to vibrate until $\frac{80}{6}$ or 13.3 sec. after the source.

Therefore, since $y = 6 \sin \pi t$ defines the motion of the source, and since in simple wave-motion each particle executes the same to and fro motion,

$$y = 6 \sin \pi \left(t - \frac{80}{6}\right)$$

defines the motion of the second particle.

Hence, when $t = 60$ sec., the displacement of the second will be given by

$$y = 6 \sin \pi \left(60 - \frac{80}{6}\right)$$
$$= 5.2 \text{ cm.}$$

This solution neglects a change of amplitude with increasing distance from the source, a question considered in Section 21.

8. To establish some fundamental relations in wave-motion, let us consider the simple case of a line of elastically connected particles, one end of which goes through a periodic motion. For simplicity, suppose that this periodic motion is similar to the to and fro motion of a simple pendulum or of a mass vibrating at the end of a spring. In Fig. 9 we have represented the position of successive particles, (1) at the instant the end particle marked O has reached its position of maximum displacement, as in *a;* (2) one-quarter of a vibration later, when the end particle is back in its normal position, as in *b;* (3) when this particle has reached its position of maximum displacement on the other side of the nor-

mal, as in *c*; (4) after one complete vibration, when particle 0 is back to its normal position, as in *d*.

The figures have been drawn on the supposition that during the time of one-quarter of a vibration, the wave-disturbance travels the distance separating six particles. In *a*, Fig. 9, therefore, particle 6 is just on the point of being displaced, while intermediate particles are displaced amounts depending on their distances from the source. In *b*, Fig. 9, particle 12 is on the point of being displaced; in *c*, Fig. 9, particle 18, while it is not difficult to see that after one

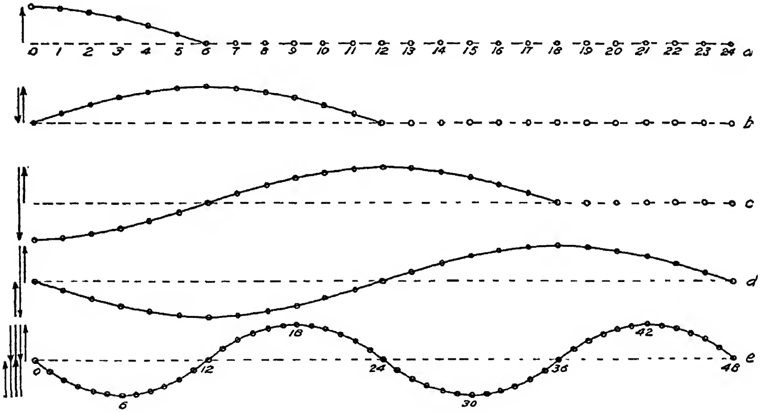


Fig. 9.

complete vibration of the source, the first 24 particles occupy positions somewhat as illustrated in *d*, Fig. 9. After two complete vibrations the configuration will be as in *e*, Fig. 9, in which case the diagram is drawn on a smaller scale.

A consideration of these diagrams will make it clear that during the time of one vibration, the wave disturbance travels a distance equal to one wave-length; during two complete vibrations, two wave-lengths and so on. In general if

λ = wave-length,
V = velocity of wave distance
T = periodic time,

it follows that, $\lambda = VT$. (2.01)

This important relation is sometimes expressed in another form by substituting for T in terms of n, the *frequency*, that is, the number of vibrations per second. Since $n = 1/T$, relation (1) becomes

$$V = n\lambda. \qquad (2.02)$$

9. The Form of a Train of Waves. — A curve passing through all the particles at any given instant represents the *wave-form*, which, therefore, in the case of a source executing the simple periodic motion we have assumed, is approximately represented in Fig. 9, diagrams *d* and *e*. If we had assumed a different kind of motion for the source, the wave-form would have been different. For example, if a complete vibration of the end particle had con-

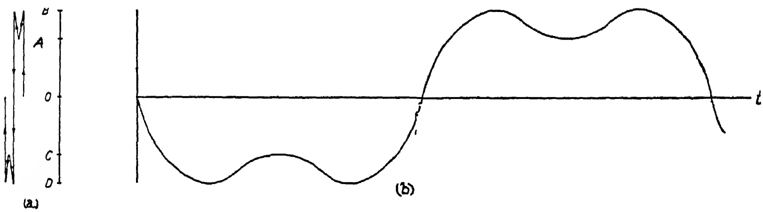


FIG. 10.

sisted of a motion from O to A to B, B back to A, A to B through O to D, D to C, C back to D, D through O, somewhat as shown in *a*, Fig. 10, the student can verify for himself that the wave-form would be as illustrated in *b*, Fig. 10. In general, the wave-form depends on the nature of the vibration of a single particle, complex wave-forms corresponding to complex periodic motions of individual particles. It will be seen presently, however, that the most complicated vibrations can be built up by combining a number of simple to and fro motions of the type we have first considered. It is highly important, therefore, to study in detail this special periodic motion, which is called Simple Harmonic (S.H.M.).

10. Simple Harmonic Motion. — First of all, let us define, with reference to Fig. 11, exactly what is meant by S.H.M. When a particle P moves around a circle with uniform velocity,

the to and fro motion of the projection of P on any diameter is called a Simple Harmonic Motion. Thus, as P moves around the circle, the point M executes a S.H.M. along the line YOY', the point N along the line XOX', and so for any diameter.* It is important for the student to guard against confusing the motion of the particle P in what is called the *circle of reference*, with the S.H.M. which is a *linear* motion. A good example of the latter is found in the vibratory motion of a mass at the end of a spring.

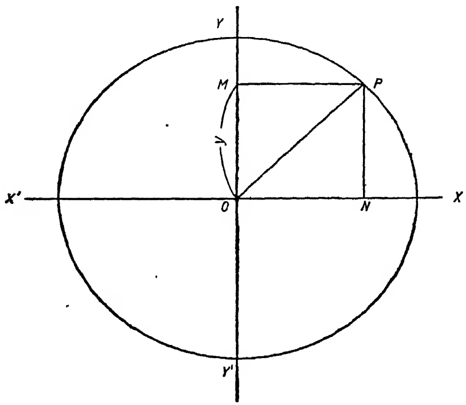


FIG. 11.

Consider now the S.H.M. of the point M along the diameter YOY'. It moves about a mean position O, having a maximum displacement, equal to OY or OY', called the *amplitude* of the motion. Suppose after t seconds it has moved from its normal position at O until its displacement $y = OM$. The corresponding particle in the circle of reference in the same time has moved from X to P, the radius having swept out the angle POX.

Then, we have,

$$\begin{aligned} y &= OP \sin OPM \\ &= a \sin POX \\ &= a \sin \omega t, \end{aligned} \tag{2.03}$$

where ω = angular velocity of the particle in the circle of reference, in radians per second. It is more usual to express the displacement in terms of T the periodic time. Since the radius sweeps

* S.H.M. may also be defined as the motion of a particle in a straight line when it moves with reference to a centre of attraction subject to the law that the acceleration is proportional to the displacement, and of opposite sign. It is left as an exercise in mechanics, for the student to prove that the two definitions mean essentially the same thing.

out a complete revolution or 2π radians in T seconds, we have at once

$$\omega = \frac{2\pi}{T},$$

and relation (2.03) becomes

$$y = a \sin 2\pi t \qquad (2.04)$$

Equation (2.04) is the definition of a S.H.M. in mathematical language. By means of it we can calculate the displacement, at any time t, of a particle executing a S.H.M. of amplitude a and period T, provided the particle is in its normal position at the beginning of the time considered. Note again that this relation has reference to a linear motion.

11. Graphical Representation of a S.H.M. — It is frequently convenient to plot a graph of equation (2.04), that is, a curve from which we can read off the displacement at any instant. This can be done in two ways.

(1) By giving t successive values, calculating corresponding values of y, and plotting y against t. In Table I this has been done for $T = 12$, $a = 10$.

Table I

t	$\frac{2\pi t}{12}$	$10 \sin \frac{2\pi t}{12}$
0	0°	0
1	30°	5
2	60°	8.6
3	90°	10
4	120°	8.6
5	150°	5
6	180°	0
7	210°	−5
8	240°	−8.6
9	270°	−10
10	300°	−8.6
11	330°	−5
12	360°	0
13	390°	5

The resulting curve is plotted in Fig. 12. It will readily be seen that, instead of using particular values of T and of a, we might give t values such as $T/12$, $2T/12$, $3T/12$, etc. In the graph, accord-

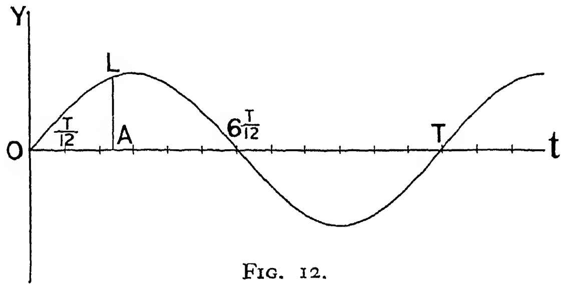


FIG. 12.

ingly, we have represented the time and the displacement in this more general way, as well as using the particular values of Table I.

(2) Since the particle in the circle of reference moves with uniform velocity, equal arcs on this circle represent equal time intervals. If we sub-divide the circle of reference into 12 equal

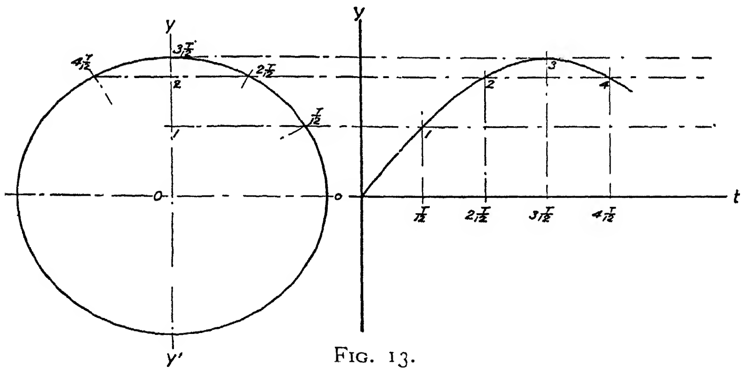


FIG. 13.

parts, for example, as in Fig. 13, we can find by dropping perpendiculars on the diameter YOY', the positions of the particle executing a S.H.M. along this line, after successive equal time intervals. Thus, if the S.H.M. particle is at O at the beginning of

the time, after $T/12$ secs. it is at 1, after $2T/12$ at 2, and so on. Hence, by taking as our time axis an extension of the perpendicular through O, and marking on it any equal distances to represent these equal time intervals, we can locate points on the curve we desire by extending the perpendiculars through the corresponding positions of the S.H.M. particle after the manner illustrated in Fig. 13. The complete curve, of course, will be exactly the same as that obtained by the other method.

The student must be careful not to confuse this curve with that representing a wave-form at any instant, as in *d* and *e*, Fig. 9. Fig. 12 is simply a graph illustrating the to and fro motion of a

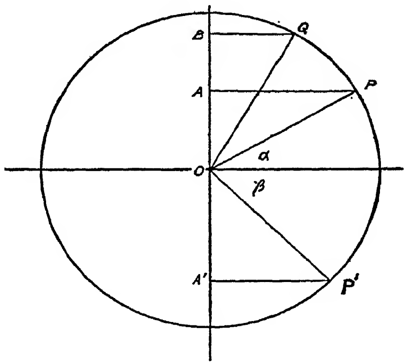


FIG. 14.

single particle. From it we may read off the displacement at any instant. For example, when the time is equal to that represented by OA, the displacement is represented by AL.

12. More General Equation of a S.H.M. — Equation (2.04) is true for a particle which is in its normal position when $t = O$. This, however, is not necessarily the case, for we may measure the time from any starting point. Suppose, for example, that when $t = o$, the S.H.M. particle is at A (Fig. 14), after t seconds at B, the corresponding positions of the particle in the circle of reference being P and Q. Then, proceeding as before, we have, at the time t,

$$
\begin{aligned}
y = OB &= a \sin BQO \\
&= a \sin QOX \\
&= a \sin (POQ + POX) \\
&= a \sin (\omega t + \alpha), \text{ where } \alpha = \angle POX \\
&= a \sin \left(\frac{2\pi t}{T} + \alpha\right.
\end{aligned} \qquad (2.05a)
$$

Similarly, if when $t = O$, the S.H.M. particle was at A', we can deduce,

$$y = a \sin \left(\frac{2\pi t}{T} - \beta\right), \text{ where } \beta = \angle XOP'. \qquad (2.05b)$$

Equations (2.05a and b), therefore, represent a little more generally than equation (2.04), a S.H.M.

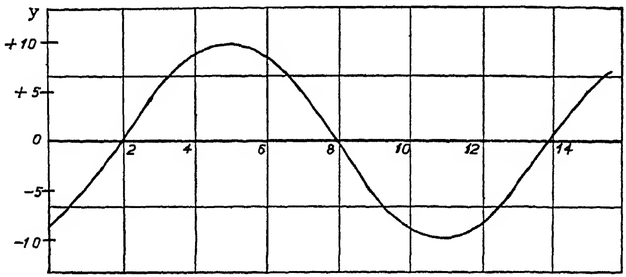


FIG. 15.

Example. — *Plot a graph representing* $y = 10\ sin\ \left(\frac{2\pi t}{12} \quad \pi\right.$

By applying either of the methods outlined above, the student can verify the solution which is given in Fig. 15.

13. Phase Angle. — From relations (2.04) and (2.05) it is at once evident that the position of a particle at any instant can be found from a knowledge of the angle,

$$\frac{2\pi t}{T} \text{ or } \left(\frac{2\pi t}{T} + \alpha\right) \text{ or } \left(\frac{2\pi t}{T} - \beta\right).$$

For that reason this angle is called the *Phase Angle*. The value of phase angle when $t = o$, is sometimes called the *Epoch*.

The following example may be of assistance to the student.

A particle executing a S.H.M. given by the relation $y = 4 \sin \left(\frac{2\pi t}{6} + \alpha\right)$ *is displaced* $+1$ *units when* $t = 0$. *Find,* (1) *the phase angle when* $t = 0$; (2) *the difference in phase between any two positions of the particle two seconds apart;* (3) *the phase angle corresponding to a displacement of* $+2$; (4) *the time to reach a displacement of* $+3$ *from the initial position.*

Solution:

(1) When $t = 0$, $y = 1$,

$$\therefore 1 = 4 \sin (0 + \alpha),$$

from which $\sin \alpha = 0.25$

and $$\alpha = 14^\circ\ 30' = .25 \text{ radian} = \frac{\pi}{12.4}$$

(2) Let t_1 and t_2 be any two times such that $t_2 - t_1 = 2$ secs.

Then since $$y_1 = 4 \sin \left(\frac{2\pi t_1}{6} + \alpha\right)$$

and $$y_2 = 4 \sin \left(\frac{2\pi t_2}{6} + \alpha\right)$$

the required phase difference

$$= \left(\frac{2\pi t_2}{6} + \alpha\right) - \left(\frac{2\pi t_1}{6} + \alpha\right)$$

$$= \frac{2\pi}{6}(t_2 - t_1)$$

$$2\pi\cdot 2 \quad = 120^\circ$$

(3) When $y = 2$, we have

$$2 = 4 \sin \left(\frac{2\pi t}{6} + \alpha\right.$$

$$= 4 \sin (\text{phase angle})$$

from which

$$\sin (\text{phase angle}) = \tfrac{1}{2}$$

or $$\text{phase angle} = 30^\circ$$

(4) When $y = 3$, we have

$$3 = 4 \sin \left(\frac{2\pi t}{6} + \alpha\right)$$

$$\therefore \tfrac{3}{4} = \sin \left(\frac{2\pi t}{6} + \frac{\pi}{12.4}\right)$$

$$\therefore \frac{2\pi t}{6} + \frac{\pi}{12.4} = \sin^{-1}$$

$$= 48^\circ \, 40'$$

from which $$t = 0.57 \text{ secs.}$$

In problems of this kind where periodic functions are involved, more than one numerical answer is possible. Thus, in part (1) of the above problem, where $\sin \alpha = 0.25$, α may not only have the value $14^\circ \, 30'$ but also $(180^\circ - 14^\circ \, 30')$, as well as $(14^\circ \, 30' + 360^\circ)$, $(14^\circ \, 30' + 720^\circ)$, etc. Note this when solving such questions as 22 and 23, page 57.

14. Composition of S.H.M. — There are many problems in light which can be solved by finding the resultant motion of a particle which is acted on simultaneously by two or more S.H.M. We shall, therefore, next discuss this general problem with special reference to two important cases: (1) when each of two or more S.H.M. if acting alone would cause *motion in the same straight line;* (2) when two S.H.M. act, the motion due to one being *in a line at right angles* to the motion due to the other.

Composition of S.H.M. in the Same Straight Line. — Two methods may be used, each of which will be considered with reference to the following problem.

Find the resultant motion when a single particle is acted on simultaneously by three S.H.M. given by the relations,

$$y_1 = a \sin \frac{2\pi t}{T}; \; y_2 \quad a \sin \frac{2\pi}{} \; t - \qquad y_3 = a \sin \frac{2\pi}{T}\left(t - \frac{}{6}\right.$$

Method I. Give t successive values; calculate for each the corresponding displacement due to each S.H.M.; find the resultant displacement by algebraic summation; and finally plot resultant values of the displacement against corresponding values of t. A few of such values are given in Table II.

Table II

t	y_1	y_2	y_3	*Resultant* y
0	0	$-0.5\,a$	$-0.86\,a$	$-1.36\,a$
$1\ T/12$	$+0.5\,a$	0	$-0.5\,a$	0
$2\ T/12$	$+0.86\,a$	$+0.5\,a$	0	$+1.36\,a$
$3\ T/12$	$+1.0\,a$	$+0.86\,a$	$+0.5\,a$	$+2.36\,a$

By obtaining in this manner a large number of corresponding values of y and t, and plotting, a curve is obtained similar to the broken one in Fig. 16. This curve then represents the required resultant motion of the particle.

The method involves an important principle, called the *Principle of Superposition*, according to which the resultant displace-

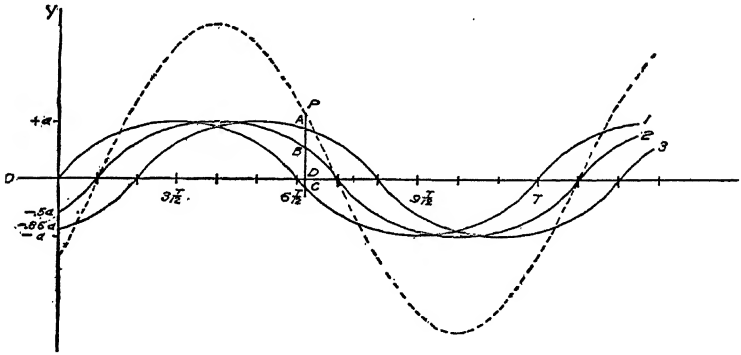


FIG. 16.

ment at any instant is equal to the algebraic sum of the individual displacements at that instant. This holds provided the displacements are small, its justification being found in the fact that it leads to correct results.

Method II. The method just outlined involves a good deal of calculation and for that reason the following more graphical method is recommended. In this the graph of each S.H.M. is plotted separately to the same scale and with the same set of coordinates. Points on the resultant graph are then obtained by

applying the principle of superposition. For example, at the time represented by OD (Fig. 16), the displacement due to the first S.H.M. is $DC(-)$, due to the second $DB(+)$, due to the third $DA(+)$, from which the resultant displacement at this instant $= DA + DB - DC = DP$. Similarly, other points are obtained in sufficient number to enable a smooth graph to be

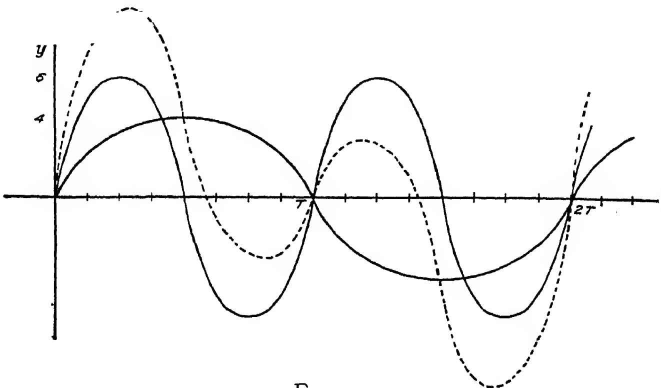


FIG. 17.

drawn. The curve obtained, of course, will be exactly the same as by Method I.

By inspection we see that the resultant motion, (*a*) is periodic; (*b*) because of its shape, almost certainly is simple harmonic; (*c*) by actual measurement has an amplitude $= 2.7\ a$.

Both these methods are perfectly general and may be applied whether the individual S.H.M. are of the same period, as in this problem, or of different periods, as in Fig. 17, which is a solution of problem 5 at the end of this chapter. The essential condition is that they all represent motion in the same straight line.

15. Resultant Amplitude for S.H.M. of the Same Period. — This case is so important (because of its application to diffraction problems) that a simple graphical method of obtaining directly the resultant amplitude will be given. Let us, first of all, consider the combination of two such S.H.M. given by,

$$y = a \sin \frac{2\pi t}{T}, \text{ and } y \quad b \sin\left(\frac{2\pi t}{T} + \alpha\right),$$

which the student should now see represent two S.H.M. of the same period, of unequal amplitude and differing in phase by α.

Suppose these two motions are executed along the line YOY^1, Fig. 18, and that at *any* time t, the displacement due to the first

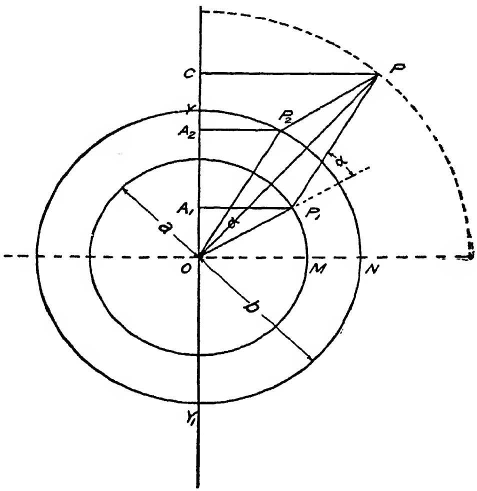


FIG. 18.

alone is OA_1, to the second alone OA_2. P_1 and P_2 will then represent the corresponding positions of the particles in the circles of reference, the angle P_1OM being equal to $\frac{2\pi t}{T}$, the angle P_2ON $= \left(\frac{2\pi t}{T} + \alpha\right)$.

Now complete the parallelogram, P_1OP_2P by drawing P_1P through P_1 parallel to the radius OP_2 and P_2P through P_2 parallel to the radius OP_1. Drop a perpendicular PC to the YOY' line.

Applying the principle of superposition, we have the resultant displacement at time $t = OA_2 + OA_1$

$$= OA_2 + A_2C,$$

since $OA_1 = A_2C$ (because OP_1 and P_2P are equal and parallel), or resultant displacement $= OC$.

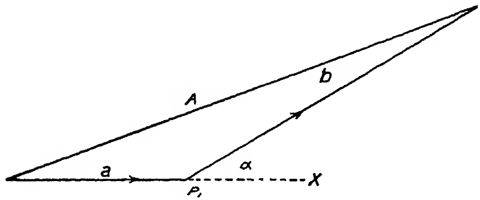


FIG. 19.

But, if the point P revolves in a circle of radius OP with the same period as the two S.H.M., then OC represents the displacement of a S.H.M. of amplitude OP, the radius of this circle, at the time P is in the position shown in the figure. Since the above construction is perfectly general, we con-

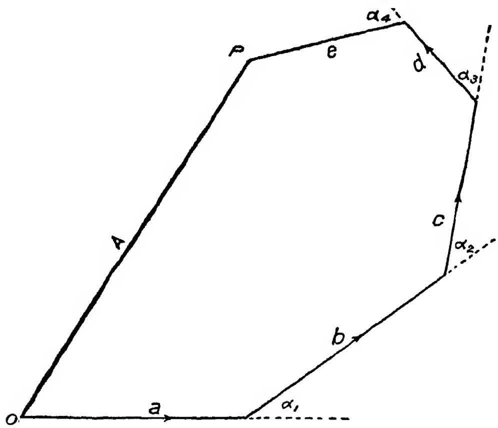


FIG. 20.

clude that the resultant of two S.H.M. of the same period is also a S.H.M. of amplitude given by OP and differing in phase from the one of amplitude OP_1 by the angle POP_1.

If, therefore, we wish only the resultant amplitude it can quickly be found, because it is represented by the side OP of the triangle OP_1P, the other two sides of which represent the two individual amplitudes. In the problem we have been discussing, for example, all that is necessary is to make a diagram like Fig. 19, in which the angle $PP_1X = \alpha$. We see, too, that since

$$OP^2 = OP_1^2 + PP_1^2 + 2OP_1 \cdot P_1P \cos \alpha,$$

the resultant amplitude A may be calculated from the relation

$$A^2 = a^2 + b^2 + 2ab \cos \alpha. \tag{2.06}$$

This method may readily be extended to any number of S.H.M. in the same straight line. For, having taken two of them, and having found the resultant OP in the manner just described, we may then combine this resultant with the third, and so on. Thus, OP in Fig. 20 represents the resultant of 5 S.H.M. of the same period, with individual amplitudes a, b, c, d, e, with phase difference between first and second $= \alpha_1$, between second and third $= \alpha_2$, etc.

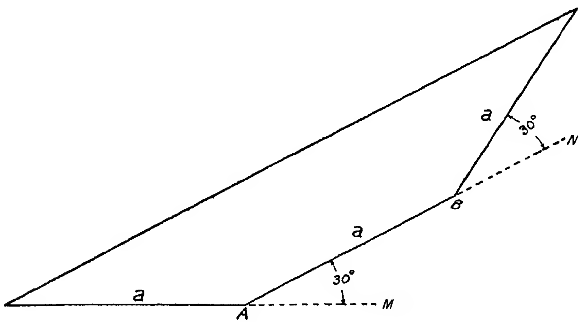


FIG 21.

An application to the problem discussed in section 14 by the general methods may perhaps help the student. Here we have 3 S.H.M. of the same period, each of amplitude a, with phase

difference between the first and second $= \frac{\pi}{6}$, between the second and third $= \frac{\pi}{6}$. The resultant amplitude is therefore represented by OC of Fig. 21. By actual measurement or by trigonometrical calculation OC is found $= 2.7a$. The same result may be obtained by actual measurement from the resultant graph of Fig. 16.

15A. Equation of Resultant S.H.M.—Referring back to Fig. 18, we see that, since the angle α, that is, P_1OP_2 represents the difference in phase between the two individual S.H.M., it follows that the angle P_1OP must represent the phase difference between the resultant S.H.M. and the one whose amplitude is OP_1. The equation of the resultant motion is represented, therefore, by

$$y = OP \sin\left(\frac{2\pi t}{T} + P_1OP\right),$$

or, referring to Fig. 19, by

$$y = A \sin\left(\frac{2\pi t}{T} + P_1OP\right).$$

Similarly the equation of the resultant of the three S.H.M. represented by Fig. 16 and Fig. 21 can be written down once Fig. 21 is constructed. It is given by

$$y = OC \sin\left(\frac{2\pi t}{T} - AOC\right)$$

$$= 2.7\, a \sin\left(\frac{2\pi t}{T} - 30^\circ\right).^*$$

If it should be required to find the equation of the resultant from Fig. 16, not from Fig. 21, then we proceed as follows.

* The minus sign is used here, because in this problem the phase difference between the first and the second S.H.M. (as well as the second and the third) is -30°. That is, in Fig. 21, the angles MAB and NBC represent minus angles, hence the angle AOC is also a minus angle. In combining S.H.M. where both $+$ and $-$ values occur, it is advisable to draw $+$ angles above the direction chosen to represent the first S.H.M. and $-$ angles below this direction.

By actual measurement (from Fig. 16) the amplitude of the dotted curve (the resultant) is found $= 2.7a$. The equation is then of the form

$$y = 2.7a \sin\left(\frac{2\pi t}{T} + \theta\right)$$

where the value of θ has yet to be found. To find its magnitude, we must use some value of y and the corresponding value of t. For example, from the graph of Fig. 16, by actual measurement,

$$y = -1.35a, \text{ when } t = 0.$$

We have, therefore,

$$-1.35a = 2.7a \sin(0 + \theta),$$

from which $\theta = -30°$.

The final equation of the resultant becomes

$$y \quad 2.7a \sin\left(\frac{2\pi t}{T} - 30^{\circ}\right.$$

Or, again, we can see from the graph that

$$y = 0, \text{ when } t = \frac{T}{12}.$$

Therefore, we have

$$0 = 2.7a \sin\left(\frac{2\pi}{T} \cdot \frac{T}{12} + \theta\right)$$

or $$\frac{\pi}{6} + \theta = 0$$

or $$\theta = -30°, \text{ as before.}$$

16. Composition of Two S.H.M. in Lines at Right Angles. — This is another example of superposition which has a very important application in the study of light. Two general methods may be used, just as in the composition of S.H.M. in the same straight line, discussed in section 14. These may best be explained by solving the following problem.

Find the resultant motion of a particle acted on by two S.H.M., in lines at right angles, subject to the relations

$$y = 6 \sin \frac{2\pi t}{T}, \quad x = 8 \sin \left(\frac{2\pi t}{2T} - \frac{\pi}{3}\right).$$

Method I. Give t a series of values, work out corresponding values of x and of y, and plot x against y. A few of such values are recorded in Table III.

Proceeding in this way, the graph shown in the upper right-hand corner of Fig. 22 is obtained, points marked 0, 1, 2, 3, etc., corresponding to the equal time intervals given in the first column of Table III. This graph shows the resultant path of the particle.

Table III

t	y	x
0	0	-6.9
$1\ T/12$	3	-5.6
$2\ T/12$	5.2	-4.0
$3\ T/12$	6	-2.1
$4\ T/12$	5.2	0

Method II. This is a more graphical method, although essentially the same as the first. It will best be understood by reference to Fig. 22. The y motion is supposed to take place along the line YOY', the x motion along $XO'X'$. By dividing the y circle of reference into, say 12 equal parts, the y displacements at the end of the corresponding time intervals, each equal to $\frac{T}{12}$, may be obtained as already explained in section 11. The same may be done for the x motion, but, since in this case the period is $2T$, compared with T for the y motion, the circle of reference must be divided into 24 equal parts to obtain time intervals of magnitude $\frac{T}{12}$. At any of these time instants, therefore, the displacement for each motion may be found graphically. Thus, when $t = 0$, $y = 0$, and the displacement for y motion is 0; when $t = 0$,

$x = 8 \sin\left(-\frac{\pi}{3}\right)$ or the phase angle $= -60°$, and the position of particle in the x circle of reference is at point P marked o, while the actual x displacement $= O'M$.

Again, when $t = 2 \cdot \frac{T}{12}$, it will be seen at once that the y displacement $= OB$, the particle in y circle of reference being at 2, while the x displacement $= O'N$, the corresponding particle in x circle of reference being also at the point 2 on that circle.

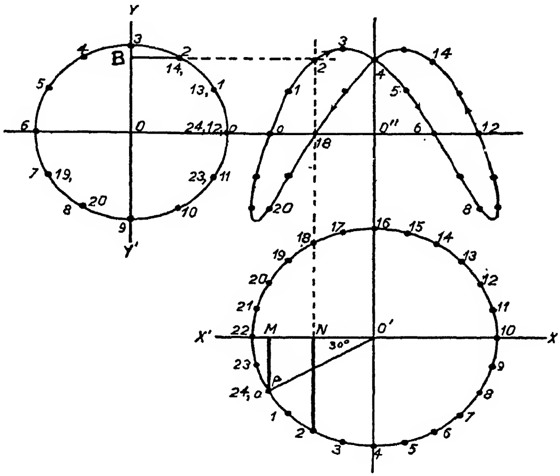


FIG. 22.

Suppose, now, we take as our origin for plotting the resultant path, the point O'', that is, the point of intersection of the perpendiculars through the normal positions for the x and the y motions. Then, by extending the perpendiculars through any other corresponding positions of x and y, the point of intersection will give the *resultant* position due to the two motions, with reference to the origin O''. Thus, the perpendicular through N (x motion) intersects that through B (y motion) at the point 2, which, therefore, is the actual position of the particle at the time $2 \cdot \frac{T}{12}$. In this

way all other points on the resultant path were obtained, after which a smooth curve was drawn through them as shown in Fig. 22. This resultant path will obviously be repeated every $2T$ seconds; that is, generally speaking, in every time interval which is an

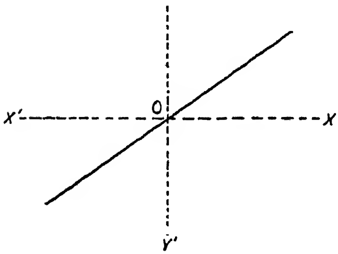


FIG. 23.

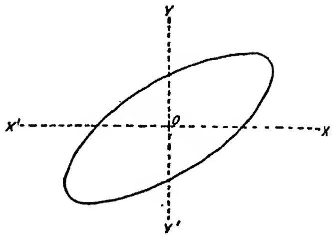


FIG. 24.

exact multiple of each of the periodic times. Of such resultant paths actual examples are provided in the well-known Lissajous figures, obtained by receiving a spot of light on a screen after reflection from two small mirrors, one at the end of the prong of

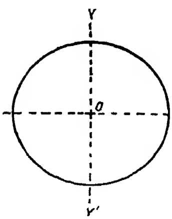


FIG. 25.

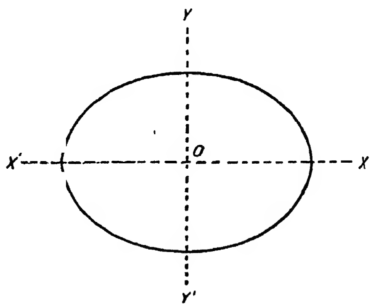


FIG. 26.

a vibrating tuning fork, the second at the end of a second fork vibrating in a line at right angles to the first.

The composition of two S.H.M. in lines at right angles is of very great importance in the study of polarized light. It is well, therefore, for the student to work out for himself the resultant of two

such S.H.M., of equal periods, amplitudes equal or unequal, with phase differences of varying amounts. In Figs. 23, 24, 25, 26, and 27, the solutions of a few cases have been given. In Fig. 23, amplitudes are unequal, phase difference = 0; in Fig. 24, amplitudes unequal,

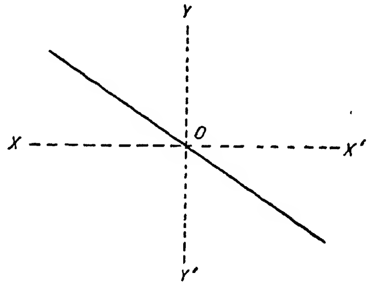


FIG. 27.

phase difference 45°; in Fig. 25, amplitudes equal, phase difference 90°; in Fig. 26, amplitudes unequal, phase difference 90°; in Fig. 27, amplitudes unequal, phase difference 180°.

17. Fourier's Theorem. — The work of section 14 shows graphically that the most complex periodic motions may be built up by combining a number of S.H.M. of varying periods and amplitudes. Conversely, we may resolve a complex periodic motion into a number of S.H.M. This is a result embodied in a mathematical theorem due to Fourier, which we may express in this way. Any function of t which is finite and continuous between values of $t = 0$ and $t = T$ may be expressed as a simple sum of sine and cosine terms. Symbolically,

$$f(t) = a_0 + a_1 \cos\frac{2\pi t}{T} + a_2 \cos\frac{4\pi t}{T} + a_3 \cos\frac{6\pi t}{T} + \cdots$$

$$+ b_1 \sin\frac{2\pi t}{T} + b_2 \sin\frac{4\pi t}{T} + b_3 \sin\frac{6\pi t}{T} +$$

where the values of the coefficients a_0, a_1, a_2, $\cdots \quad \cdots$ b_1, b_2, $\cdots \quad \cdots$ can be calculated from the form of the function.

Since $y = b_n \sin \dfrac{2n\pi t}{T}$ and $y = a_n \cos \dfrac{2n\pi t}{T}$

$$= a_n \sin \left(\frac{2n\pi t}{T} - \frac{\pi}{2}\right)$$

represent S.H.M., our graphical method of adding a number of S.H.M. in the same straight line, is seen to be just a special example of Fourier's theorem.

18. Equation of a Plane Wave. In sections 10 to 17 we have been discussing the periodic motion of a single particle. We return now to our consideration of wave-motion. If a S.H.M. gives rise to a train of waves in a medium, what relation describes the resulting wave-motion?

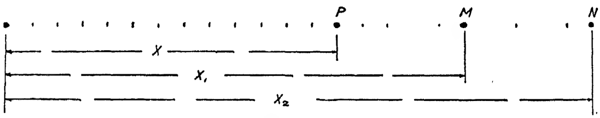


FIG. 28.

To answer that question, consider first a line of particles along which waves are travelling at the rate of V cm. per sec., and define the position of any particle P by its distance x from the source O. Suppose, further, that the source executes a S.H.M. given by the relation

$$y = a \sin \frac{2\pi t}{T}.$$

Then, since each particle executes the same motion, and since the particle at P does not begin to vibrate until $\dfrac{x}{V}$ seconds after that at O, it follows that the motion of P is given by

$$y = a \sin \frac{2\pi}{T}\left(t - \frac{x}{V}\right) \qquad (2.07a)$$

or since $\lambda = VT,$

by $$y = a \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right). \qquad (2.07b)$$

Equations (2.07a) and (2.07b) enable one to find the displacement of any particle at any time, (2.07a) in terms of a, T, and V; (2.07b) in terms of a, T, and λ. They are, therefore, two forms of the fundamental equation of a plane wave travelling along the x axis. In either of them we have combined into a single statement the two characteristics of wave-motion discussed at the beginning of this chapter.

19. Difference in Phase between any two Particles in a Wave-Train — Phase Difference and Path Difference. — Consider any two particles such as x_1 and x_2 (Fig. 28), in a line along which a train of simple waves is travelling. The motion of x_1 is given by

$$y_1 = a \sin \left\{ 2\pi\left(\frac{t}{T} - \frac{x_1}{\lambda}\right) \pm \alpha \right.$$

and of x_2 by

$$y_2 = a \sin \left. 2\pi\left(\frac{t}{T} - \frac{x_2}{\lambda}\right) \pm \alpha \right\}.$$

Therefore, at any instant, the difference in phase for these two particles

$$2\pi\left(\frac{t}{T} - \frac{x_1}{\lambda}\right) \pm \alpha \Big\} - 2\pi\left(\frac{t}{T} - \frac{x_2}{\lambda}\right) \pm \alpha$$

$$\frac{2\pi}{\lambda}\left(x_2 - x_1\right)$$

$$\frac{2\pi}{\lambda} \text{ path difference.} \qquad (2.08)$$

This is a very important relation of which use will be made later. It is evidently applicable to express the difference in phase between two disturbances coming to a single particle, such as P, Fig. 29, from two sources such as S_1 and S_2, provided the two sources execute the same S.H.M.

20. Plane and Spherical Waves. — In section 18 we considered only a line of particles. Generally we have to do with a disturbance travelling through a medium of three dimensions. In that case we must distinguish between plane and spherical wave-fronts.

A *plane* wave-front is a plane surface through which a wave-disturbance is passing and in which all vibrating particles are in the same phase. If, for example, the reader imagines a large number of lines of particles, all parallel to the one of Fig. 28, and all ending in points similar to O, which lie on a vibrating plane,

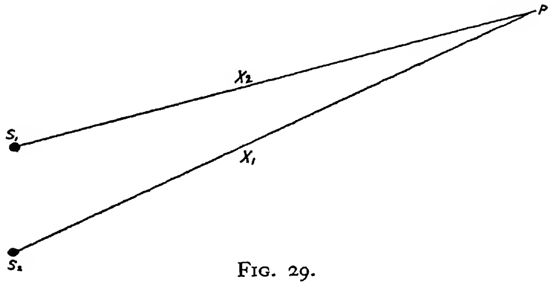


FIG. 29.

a rough picture will be obtained of plane waves. In practice, plane waves are either a special case of, or derived from, spherical waves.

A *spherical* wave-front is a portion of a spherical surface through which a wave-disturbance is travelling, and in which all vibrating particles are in the same phase.

Consider a vibrating source O, Fig. 30, so placed that a wave-disturbance travels outward in a medium in *all* directions. If the medium is homogeneous, the wave-velocity will be the same in every direction, and a portion of any sphere with centre O will represent a spherical wave-front. If the sphere is sufficiently large and a relatively small portion such as MN is taken, this part will be approximately plane. When, therefore, we have to do with such a small part of a spherical surface, we may with little error treat it as a plane wave-front. For example, if we consider the light coming to an observer from a small region on the sun's surface, we are justified in treating the wave-fronts as plane.

It will be seen from their derivation that equations (2.07a) and (2.07b) are applicable to plane waves travelling in a direction

parallel to the x-axis. They are also applicable to spherical waves travelling parallel to the x-axis, provided the fact that the amplitude changes is taken into consideration. How this can be done will be shown by a consideration of energy changes.

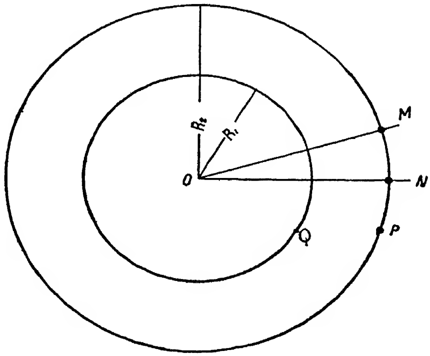


FIG. 30.

21. Change of Amplitude with Distance for Spherical Waves. — If reference is again made to Fig. 30, it will be evident that (provided there is no absorption by the medium) the same amount of energy per second traverses the sphere of radius R_1 as the sphere of radius R_2.

Let E = energy radiated per second by the source; then $\frac{E}{4\pi R_1^2}$ = energy passing each second through each sq. cm. of the sphere of radius R_1, and $\frac{E}{4\pi R_2^2}$ = energy per second per sq. cm. for sphere of radius R_2. But the intensity of the disturbance (the brightness in light) at any point such as P depends on the rate at which a particle or particles in the neighborhood of P are receiving energy. Hence,

$$\frac{\text{intensity at } Q}{\text{intensity at } P} = \frac{\frac{E}{4\pi R_1^2}}{\frac{E}{4\pi R_2^2}} = \frac{R_2^2}{R_1^2},$$

from which we deduce that, in the case of spherical waves, the intensity falls off inversely as the square of the distance from the source. The student will recognize this as the law used in ordinary photometry for comparing the candle powers of different sources.

To find out how the amplitude varies with the distance, it is necessary to know how the energy of a vibrating particle depends on its amplitude. That is essentially a problem in mechanics and it is left to the student to prove that the energy is proportional to the square of the amplitude.

Since, then, (1) the energy received by a particle per second from a distant source falls off inversely as the square of the distance from that source, and (2) the energy of a vibrating particle is proportional to the square of its amplitude, it follows that the amplitude falls off inversely as the first power of the distance. For spherical waves, we may write, therefore,

$$y = \frac{A}{x} \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right), \tag{2.09}$$

where A = amplitude at unit distance.

Relations (2.07a), (2.07b), and (2.09) are applicable only in the case of perfectly transparent media for which there is no falling off in intensity due to absorption. If absorption is to be considered, then the amplitude will gradually decrease on that account according to an exponential law, and relation (2.07b), for example, will now become

$$y = ae^{-\beta x} \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right), \tag{2.10}$$

where β may be defined as a coefficient of absorption.

22. Shape of a Train of Waves. — In section 9 it was shown generally that the shape of a train of waves depended on the nature of the periodic motion of the source (or any other particle). We are now in a position to find the relation defining the shape when the periodic motion is S.H.M. Our problem is to express y, the displacement, as a function of x, at any instant, or, in general terms, to find out how the displacement of successive particles

varies with their position *at any given moment.* In the case of water waves, we might do this by taking an instantaneous photograph. Generally, if we have plane waves defined by

$$y = a \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right),$$

we can find their shape by giving t any value we please. Take $t = T$, for example, then *at that instant,*

$$\begin{aligned} y &= a \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) \\ &= a \sin\left(2\pi - \frac{2\pi x}{\lambda}\right) \\ &= -\,a \sin \cdot^{2\pi x} \qquad (2.11) \end{aligned}$$

Figure 31 is a graph of relation (2.11). Note that while the curve has exactly the same form (although reversed) as that of

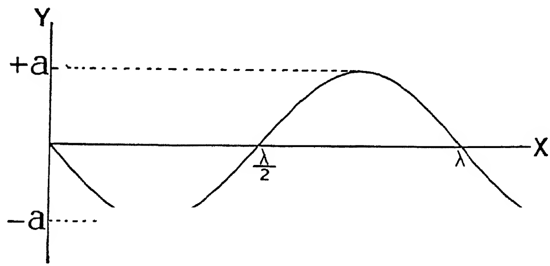


FIG. 31.

Fig. 12, it represents something entirely different. From Fig. 31 we may read off the displacements of *successive particles* at the same instant, while Fig. 12 enables us to find the displacement of a single vibrating particle at any time.

23. Composition of Wave-Trains · Stationary Waves.— When a medium is traversed by more than one set of waves, the principle of superposition may be applied to find the resultant

wave-form, in much the same way as for a single particle. An important example is found in the phenomenon of stationary waves.

Suppose a train of waves is meeting an exactly similar one travelling in the opposite direction (a condition which may readily be realized by reflection). At some instant which we may conveniently take as our initial time, the wave trains (if acting alone) will be represented as in *A*, Fig. 32. The continuous line represents a set moving to the right, the broken line the set moving

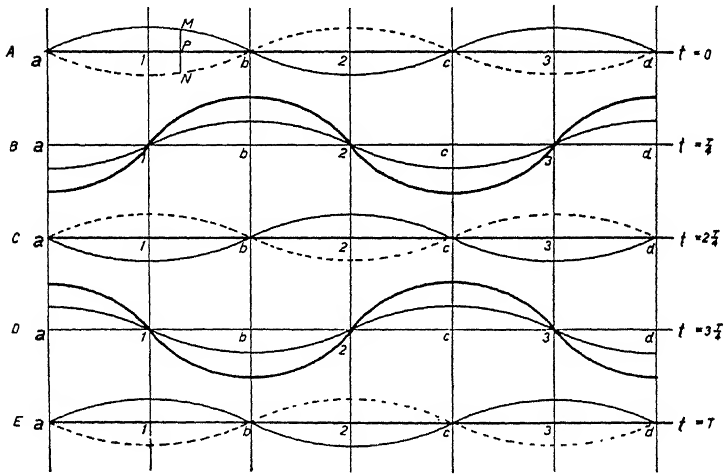


FIG. 32.

to the left. After a time equal to one-quarter of a period, each wave-train will have advanced one-quarter of a wave-length $\left(\frac{\lambda}{4}\right)$, and the position of each train will be as in *B*, Fig. 32. It will be noted that in the diagrams the distances $a1 = 1b = b2 = 2c = c3 = 3d = \frac{\lambda}{4}$. Similarly, after further time intervals of a quarter of a period, the wave-train will occupy the positions shown in diagrams *C*, *D*, and *E*. Now let us apply the principle of super-

position. In A a particle normally at P would be displaced an amount PM above its normal position due to one wave-motion, while due to the other, the displacement would be an equal amount PN in the opposite direction. The resultant displacement of P at this instant is, therefore, zero. Proceeding in this way with all the particles in each diagram, it is not difficult to see that the resultant wave-forms will be as shown by the heavy lines in the diagrams.

The resultant motion is characterized by the following important features. (1) At certain instants, as in A, C, and E, all particles are in their normal positions. (2) Certain particles such as those marked 1, 2, and 3, exactly half a wave-length apart, *are always at rest.* They are said to be at *nodal points* or *nodes.* (3) Certain other particles such as a, b, c, and d, also half a wave-length apart, execute a periodic motion with double the amplitude caused by either of the wave-trains acting alone. These particles are at *anti-nodes* or *loops.*

Waves with these characteristics are called stationary waves.

24. Analytical Proof of Nodes and Loops. — In section 18 we saw that a train of simple waves travelling in a direction parallel to the x — axis is defined by the relation

$$y = a \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right).$$

This relation was derived on the supposition that the wave-disturbance was advancing in the direction of increasing values of x. If we have an exactly similar set travelling in the opposite direction, it will be defined by

$$y = a \sin 2\pi\left(\frac{t}{T} + \frac{x}{\lambda}\right).$$

If two such sets are traversing a medium simultaneously, then the resultant disturbance will be given by

$$\begin{aligned} y &= a \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) + a \sin 2\pi\left(\frac{t}{T} + \frac{x}{\lambda}\right) \\ &= 2a \sin \frac{2\pi t}{T} \cos \frac{2\pi x}{\lambda}, \end{aligned}$$

by trigonometrical expansion.

When $x = \frac{\lambda}{4}, 3\frac{\lambda}{4}, 5\frac{\lambda}{4}, \ldots$

$$y = 2a \sin \frac{2\pi t}{T} \cos \frac{n\pi}{2}, \text{ where } n = 1, 3, 5, \ldots$$
$$= 0, \textit{for all values of } t.$$

In other words, there are certain particles half a wave-length apart, which are always at rest. These are at nodal points.

Similarly, when $x = 2\frac{\lambda}{4}, 4\frac{\lambda}{4}, \ldots$

$$y = 2a \sin \frac{2\pi t}{T} \cos m\pi, \text{ where } m = 1, 2, 3, \ldots$$
$$\pm 2a \sin \frac{2\pi t}{T}.$$

That is, there are certain other particles, also half a wave-length apart, and mid-way between the nodal points, which execute a S.H.M. with double the amplitude of that due to either of the original disturbances, when acting alone. These particles are at loops.

Problems

(In all questions, unless otherwise stated, assume that waves are plane.)

1. Two stakes stand $5m$. apart in a pond. A stone is dropped into the water in line with the stakes. At a certain instant it is observed that there is a crest at each stake and five crests between them. What is the wave-length of the disturbance?

2. A system of water waves is moving across a pond, $\lambda = 200$ cm. At a certain instant a crest is at a stake standing in the water. A, B, C, and D are points between the stake and the source of the waves. Their distances from the stake are 2, 5, 8.5, and 11 meters respectively. Is the surface of the water above or below the normal level at these points? Is it moving up or down?

3. Plot the curve $y = 6 \sin \frac{2\pi x}{3}$.

4. Two particles execute motions subject to the laws $y = 8 \sin \frac{2\pi t}{10}$ and $y = 8 \sin \left(\frac{2\pi t}{10} + \frac{\pi}{3}\right)$.

Express the position of each particle with reference to its normal position, (*a*) at time $t = 0$, (*b*) at time $t = 5$.

5. A particle is acted on simultaneously by two S.H.M. given by

$$4 \sin \frac{2\pi t}{6} \text{ and } y \quad 6 \sin \frac{2\pi t}{3}$$

Plot a graph showing the resultant motion.

6. A particle is acted on by two simple harmonic motions of amplitudes 6 and 8 units, of equal periods, differing in phase by 60°. Find graphically the resultant amplitude.

7. A particle vibrates under the influence of three S.H.M. (of the same period, all in the same direction) with amplitudes 3, 4, and 5; with phase differences, 30° between the first and second, 45° between the second and third. Find the resultant amplitude of the particle.

8. A particle is subjected simultaneously to 8 simple harmonic motions, all of the same period, all of equal amplitude. If the difference of phase between each successive two is the same, find what this phase difference must be, so that the resultant amplitude of the particle is zero.

9. A particle is subjected simultaneously to 2 S.H.M. executed in lines at right angles. Find graphically the resultant motion when the individual S.H.M. are given by:

$$(a)\ y = a \sin \frac{2\pi t}{T};\ x = b \sin \left(\frac{2\pi t}{T} - \frac{\pi}{2}\right).$$

$$(b)\ y = a \sin \frac{2\pi t}{T};\ x = b \sin \frac{2\pi t}{T}.$$

$$(c)\ y = a \sin \frac{2\pi t}{T};\ x = b \sin \frac{2\pi t}{2T}.$$

$$(d)\ y = a \sin \frac{2\pi t}{T};\ x = b \sin \left(\frac{2\pi t}{2T}\right.$$

$$(e)\ y = a \sin \frac{2\pi t}{T};\ x \quad b \sin \frac{2\pi t}{T}$$

10. Plot a graph representing the S.H.M. $y = 10 \sin \left(\frac{2\pi t}{12} - \frac{\pi}{\ }\right.$

11. When a particle is acted on by two S.H.M. and circular oscillations result, write down a suitable equation for each S.H.M.

12. Between two points A and B in a certain line of particles along which a train of simple waves is travelling, there are $10\frac{1}{3}$ waves. What is the difference in phase between the particles at A and B?

13. Two particles A and B begin executing S.H.M. at the same instant, in the same phase, in the same direction, with amplitudes respectively 6 and 8, and periods 1 and 2. The waves sent out from each disturb a third particle C, 20 cm. from A and 60 cm. from B. If the

velocity of the wave motion is 40 cm. per sec., find the resultant displacement of C, 5 seconds after the particles have started to vibrate.

14. Two particles, A and B, execute S.H.M. given by the relations, $y = 6 \sin 5\pi t$ and $y = 12 \sin 10\pi t$.

If they send out waves in a continuous medium, travelling at the rate of 400 cm. per second, find the resultant displacement of a particle 200 cm. from A and 300 cm. from B, after 2 seconds.

15. A particle at the end of a long line of particles executes a periodic motion determined by the relation $y = 6 \sin 5\pi t$, as a result of which waves are sent along the line. If the velocity of the waves = 400 cm. per sec. find the displacement of a particle 20 cm. from the origin, 2 sec. after the particle at the origin began to move from its position of rest.

16. Two particles A and B executing disturbances given by the relations $y = 6 \sin 2\pi t$ and $y = 8 \sin \pi t$, send out waves which travel at the rate of 200 cm. per second and disturb a third particle C distant 200 cm. from A and 400 cm. from B.

Draw a rough graph to represent the complete motion of C starting from the instant it first began to vibrate.

17. A particle executing a periodic motion subject to the relation $y = 6 \sin \dfrac{\pi t}{6}$ sends out waves which travel through a homogeneous medium at the rate of 6 cm. per second.

(*a*) What will be the difference in phase between any two particles 12 cm. apart on any radius drawn from the vibrating source?

(*b*) What will be the displacement of a particle 80 cm. from the source one minute after the source began to vibrate?

18. Two particles A and B executing S.H.M. given by the relations $y = 6 \sin 2\pi t$ and $x = 6 \sin \pi t$, send out plane polarized waves in planes at right angles. Determine graphically the resultant motion of a particle C, 100 cm. from A and 200 cm. from B.

Velocity of wave disturbance = 400 cm. per sec.

19. Compound three trains of waves of lengths in the ratio of 1, $\frac{1}{3}$, and $\frac{1}{5}$ and of amplitudes 3, 2, and 1, starting in the same phase.

20. Compound two trains of waves of lengths 5 and 4 and of equal amplitudes.

21. When a train of plane waves of wave-length 240 cm. traverses a medium, individual particles execute a periodic motion given by the relation,

$$y = 4 \sin \left(\frac{2\pi t}{6} + \alpha\right)$$

(*a*) Find the velocity of the wave disturbance.

(*b*) Find the difference in phase for two positions *of the same particle* which are occupied at time intervals 1 second apart.

(*c*) What is the difference in phase, *at any given instant*, of 2 particles 210 cm. apart?

(*d*) If the displacement of a certain particle at a certain time is 3 cm., find where it will be 2 seconds later.

22. A particle executing a S.H.M. given by the relation $y = 6 \sin \left(\frac{2\pi t}{-} - \alpha\right)$ is displaced 3 units on the negative side of its normal position at the beginning of the time. Find:

(*a*) how long it will take to reach a position where the displacement is $+1$ unit;

(*b*) the phase angle when $t = 2$;

(*c*) if waves are sent out in a surrounding medium at the rate of 40 cm. per sec., the displacement of a second particle 20 cm. from the first when $t = 6$ sec.

23. A source particle executing a S.H.M. given by $y = 5 \sin \pi t$ sends out waves which travel at the rate of 45 cm. per second. At a certain instant the source is displaced $+2.5$ units, while a second particle x cm. from it, is at the same instant displaced $+0.868$ units. Find the value of x.

24. A particle is acted on simultaneously by two S.H.M., $y = 8 \sin \frac{2\pi t}{}$ and $y \quad 6 \sin \frac{2\pi t}{12}$.

(*a*) Find graphically, (i) the period of the resultant; (ii) the amplitude of the resultant; (iii) the position of the particle 9 seconds after the beginning of its motion.

(*b*) State whether the resultant motion is S.H.M. or not, giving reasons for your answer.

(*c*) If the particle sends out waves which travel at the rate of 400 cm. per sec., find the wave-length.

25. A particle is acted on simultaneously by the following three S.H.M.:

$$= 3 \sin \frac{2\pi t}{T}; \; y = 4 \sin \; \frac{2\pi t}{T} - 45' \quad y = 5 \sin \; \frac{2\pi t}{} + 30^\circ\Big).$$

(*a*) Find the amplitude of the resultant motion.

(*b*) Write down the equation of the resultant motion.

26. A plane wave travelling along the x-axis in the positive direction is defined by

$$y = 6 \sin (3.14t - 0.314x),$$

where the phase angle is expressed in radians when t is in seconds and x in centimetres. Find: (i) the periodic time; (ii) the wave-length; (iii) the velocity of the wave; (iv) the phase difference (in degrees) between two particles 7 cm. apart, at the same instant.

27. (*a*) At a certain instant a particle executing a S.H.M. of amplitude 6 units and period 10 units, has a phase angle of -30°. Find (i) its

phase angle; (ii) its displacement; (iii) its direction of motion with reference to its normal position, 6 units of time after this instant.

(*b*) If this particle A is the source of a wave disturbance travelling through a medium at 50 units of length per unit time, find the phase angle of a second particle 210 units distant from A (in the direction toward which the waves are travelling) at the particular moment that A, having started from its normal position, has completed 1000.1 oscillations.

28. When plane waves of length 240 cm. traverse a medium, at a certain instant a particle A has a displacement of $+ 4.0$ cm. and the displacement is increasing. Find (i) the amplitude of the disturbance if at this instant the phase angle of A is $45°$; (ii) the displacement and phase angle of a particle B 20 cm. farther away from the source than A, at the same moment.

29. Plane waves travel along the x-axis in the positive direction at a velocity of 50 cm. per second from a source at the origin moving subject to the law

$$y = 10 \sin \frac{2\pi t}{12}.$$

Find: (*a*) the wave-length; (*b*) the equation giving the shape of the wave-train when $t = 15$; (*c*) the phase angle of a particle at $x = 15$, at time $= 15$; (*d*) the phase difference at any instant between 2 particles 20 cm. apart; (*e*) after what time the equation of the shape is $y = 10 \sin \frac{2\pi x}{\lambda}$, where λ is the wave-length.

30. Two S.H.M. each of periodic time 6 units, in the same straight line, have such a phase difference that when they are superimposed the resultant displacement is $- 1.35$ units when $t = 0$, and $+ 1.35$ when $t = 1$. Find the equation of the resultant motion.

31. (*a*) Prove that a particle remains undisturbed when it is acted on by 6 S.H.M., in the same straight line, of equal amplitudes, equal periods, but with phase difference between each successive pair $= 60°$.

(*b*) If this phase difference is $90°$, find the resultant amplitude in terms of the amplitude of one of the S.H.M.

32. A plane wave arising from a S.H.M. is travelling along the x-axis in the positive direction. At a certain instant, the phase angle of the particle at the origin is $80°$, that of a particle P 20 cm. away is $20°$. If the displacement (at this instant) of the origin is 4 cm. find (i) the wave-length; (ii) the displacement of P at this instant; (iii) the equation giving the shape of the wave at this instant.

33. At a certain instant the shape of a train of waves is given by

$$y = 6 \sin \left(\frac{2\pi x}{120} - 30^c\right.$$

(*a*) Plot the curve giving the shape.
(*b*) Write down the wave-length.

(*c*) Write down the phase angle of particles at $x = 0$, $x = 10$, and $x = 40$.

(*d*) Assuming that the source (at $x = 0$) follows the law $y = 6 \sin \pi t$, find two values of t for which the above equation will represent the shape.

34. S.H.M. plane waves of amplitude 12 units and periodic time 10 secs. travel along a line of particles at the rate of 40 cm. per second. At a certain instant a particle A has a displacement of $+6$ *decreasing*.

(*a*) Find the smallest value of the phase angle of A at this instant.

(*b*) *At the same instant*, find (i) the phase angle, (ii) the displacement of a particle B 100 cm. *nearer* the source than A.

(*c*) At the same instant, plot at least one complete wave-length in a region which includes the particle A.

(*d*) Find the phase angle of A $\frac{5}{6}$ of a second *after* the "certain instant."

35. Plane waves of amplitude 8 units and periodic time 2 units travel along the x-axis with a velocity of 20 units of length per unit of time. At a certain instant, the phase angle of a particle A is 60°. Find, at this instant, (i) the displacement of A, (ii) its direction of motion (towards or away from the normal position), (iii) the phase angle of a particle B, 10 units of length beyond A, and (iv) three possible values for the position of a particle whose phase angle is 0°.

CHAPTER III

REFLECTION AND REFRACTION ON HUYGENS' PRINCIPLE

Before reading this chapter, the student who is not thoroughly familiar with the simple laws of reflection and refraction, as well as with the geometrical means of locating images in mirrors and in lenses by the use of principal rays, is strongly advised to review this work with the help of any good elementary text.

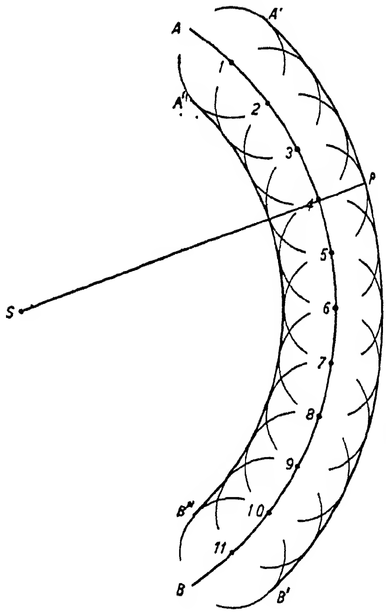


FIG. 33.

25. In Chapter I it was stated that a simple explanation of the laws of reflection and of refraction could be given from the view-point of waves. To do this we shall make use of what is called *Huygens' Principle.* Huygens, it will be recalled, was one of the supporters of the wave-theory at a time long before the phenomenon of interference had been discovered. Indeed he is generally considered to be the founder of this theory. It was necessary for him not only to deduce the laws of reflection by means of waves, but also to account for approximate rectilinear propagation. To some extent the solution of these questions is found in

this principle. According to it, if a wave-front originating in some source S, Fig. 33, has reached the position AB, each vibrating particle in this wave-front must be considered as a secondary source from which spherical wavelets spread out. After a time t, therefore, there will be a series of "wavelets" with centres 1, 2, 3, 4 . . . each of radius Vt, where V is the velocity of the

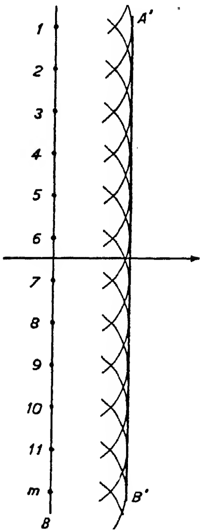


FIG. 34.

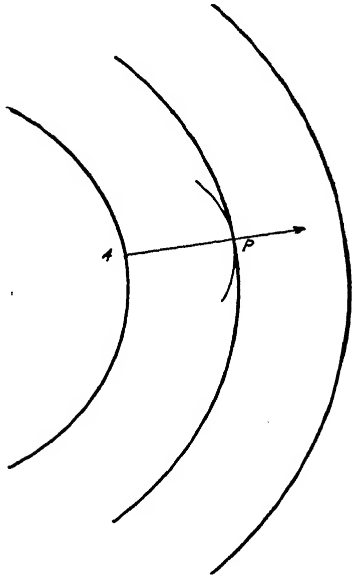


FIG. 35.

wave-disturbance. The principle states further that the surface which envelops all these wavelets constitutes a new wave-front. It will be evident from Fig. 33 and Fig. 34 that spherical wave-fronts will be propagated as spherical, plane wave-fronts as plane.

The principle as thus stated tacitly assumes that a wavelet is effective only at the small portion which is tangent to the

enveloping surface, and defines a ray as the line joining the centre of a wavelet, such as 4, Fig. 35, to this small portion at P. While in a way it thus "explains" rectilinear propagation, the principle is open to one or two objections.

(1) Why is a wave-front not propagated backwards, giving rise to $A''B''$ (Fig. 33) as well as to $A'B'$?

(2) What right has one to neglect the large portion of the wavelet not tangent to the enveloping surface? For example, in Fig. 34, why is there no edge effect from the wavelet originating in particle m?

Objection (1) may be met by taking into account the fact that particles such as 1, 2, 3, Fig. 33, are not independent sources, but are all disturbed as the result of a wave-motion coming from the source S. When this is taken into consideration it is not difficult to explain the absence of a back wave-front. Objection (2) raises the whole question of diffraction and rectilinear propagation, phenomena which Huygens' Principle, as stated above, without modification is not able satisfactorily to explain. Later we shall see how, by assuming that a wavelet is effective over the whole front hemisphere, and by using the principle of interference, Fresnel was able to extend Huygens' Principle and to give an adequate explanation of all such phenomena. In this chapter we shall deal with cases where diffraction effects may be neglected and show how, from Huygens' Principle in its original form, we can deduce the common laws of reflection and refraction.

26. Plane Waves Striking a Plane Reflecting Surface Obliquely. — Suppose a set of plane waves is incident in the direction of the arrow I, Fig. 36, on the plane reflecting surface XY. Let OM represent the position of a wave-front at a certain instant. After a time, had there been no change of medium, the wave-front would have advanced to the position DO'. Actually, therefore, particles in the surface XY between O and D will be *successively* disturbed as the wave-front advances, and each particle becomes the source of a reflected wavelet. In the figure the

positions of wavelets originating from particles O, A, B, C are drawn at the instant the particle D is on the point of being disturbed. The wavelet from O, for example, has radius $= OO'$; that from $A = AA'$, and so on for the others.

Now, applying Huygens' Principle we see that these wavelets give rise to a plane reflected wave-front DK, which will continue

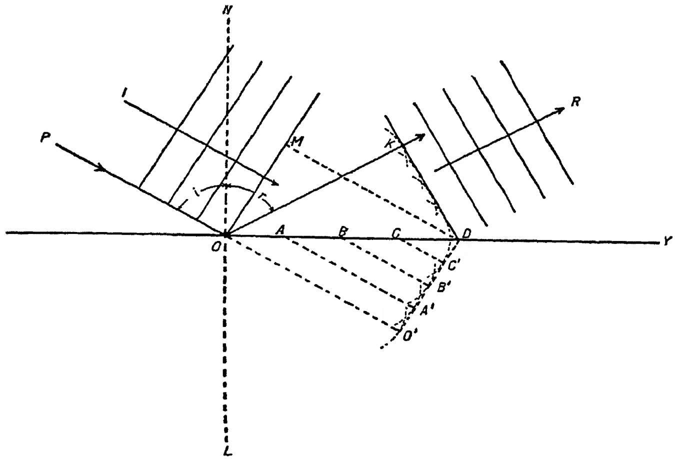


Fig. 36.

to propagate itself in the direction of the arrow R. By taking an incident *ray* such as PO and the corresponding reflected ray OK, and drawing a normal to the surface at O, it is not difficult to prove by elementary geometry that the angle NOP = angle NOK; that is, that the angle of incidence = the angle of reflection.

27. Plane Waves Striking a Refracting Surface Obliquely. — Suppose a plane wave-front OM, Fig. 37, is incident on the plane refracting surface XY separating two transparent media in which the velocities of the wave-disturbance are v_1 and v_2, $v_1 > v_2$. As in the previous case particles in the surface are successively dis-

turbed, and in the figure we have represented the position of several wavelets at the instant the edge M of the original wave-front has reached the point D. Putting it otherwise, the wavelets are drawn at the time that OM would have reached $O'D$ had there been no change in medium. The radius OK of wavelet

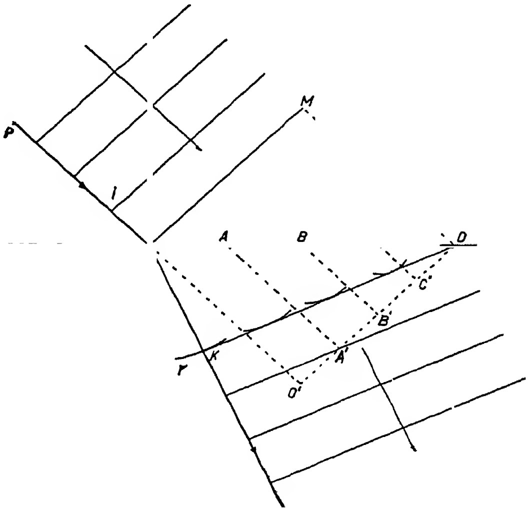


FIG. 37.

originating in O will, therefore, $= OO' \cdot \frac{v_2}{v_1}$; the radius of that from $A = AA' \cdot \frac{v_2}{v_1}$; and so on for other points. Again we see the formation of a plane refracted wave-front DK, travelling in a direction making a smaller angle with the normal to the refracting surface than the incident train.

Deduction of the Sine Law. — Consider an incident ray PO (Fig. 37) and the corresponding refracted ray OK.

If i = angle of incidence, and r = angle of refraction, then

$$\frac{\sin i}{\sin r} = \frac{\sin PON}{\sin KOL} = \frac{\sin MOD}{\sin KDO} = \frac{v_1}{v_2}.$$

By definition the index of refraction of a medium is the ratio of the velocity of light c in a vacuum * to the velocity in the given medium. Hence, introducing c into the above relation, we may write

$$\frac{\sin i}{\sin r} = \frac{\dfrac{c}{v_2}}{\dfrac{c}{v_1}} = \frac{n_2}{n_1},$$

or

$$n_1 \sin i = n_2 \sin r,$$

where n_1 is the index of the medium in which the velocity is v_1 and n_2 bears a similar relation to v_2.

If, following the notation often used in geometrical optics, we put $i = \theta_1$ and $r = \theta_2$, we obtain the convenient symmetrical form

$$n_1 \sin \theta_1 = n_2 \sin \theta_2. \qquad (3.00)$$

28. Total Reflection. — According to Huygens' Principle no special hypothesis is necessary to explain simultaneous reflection and refraction because if particles such as O, A, B, C, in either Fig. 36 or Fig. 37, are disturbed, both reflected and refracted wavelets are to be expected. Our problem is rather to explain why, under certain conditions, the refracted wave-front is absent.

The student will recall the phenomenon of total reflection. If a ray travelling in a more dense medium strikes a plane refracting surface sufficiently obliquely, there is no refraction and total reflection occurs. We shall examine, therefore, what explanation can be given of this fact by means of Huygens' Principle. First of all, the student should note that to obtain a refracted wave-

* In handbooks and tables of constants indices are given with respect to air at 760 mm. pressure and a standard temperature.

front, it is not necessary to draw a large number of wavelets. In either of the cases illustrated in Figs. 36 and 37, all that is necessary is to draw through D the plane DK which is tangent to the *single* wavelet originating in O. If then we wish the refracted

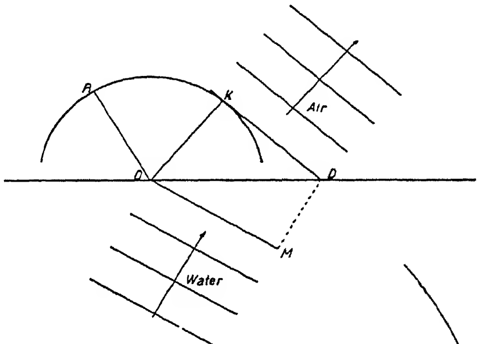


FIG. 38.

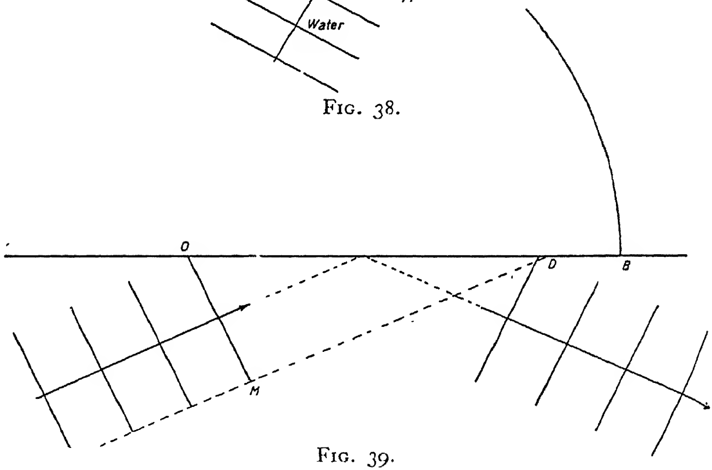


FIG. 39.

wave-front corresponding to a plane wave-front OM, Fig. 38, striking a plane surface separating water and air, at nearly normal incidence, we can proceed by, (1) drawing a wavelet with centre O and radius $OR = 1.33MD$, and (2) drawing through D the plane DK tangent to this wavelet.

Suppose, however, the incident wave-front strikes the surface very obliquely, as in Fig. 39. In that case the radius (OB) of the

wavelet originating at O is so large that it is impossible to draw a plane through D tangent to it. A refracted wave-front is then impossible and we have total reflection. Obviously, there will be a certain critical obliquity lying between the two cases illustrated in Figs. 38 and 39.

This limiting case will occur when the wavelet originating in O just reaches D, that is, when its radius $= OD$, as in Fig. 40. In

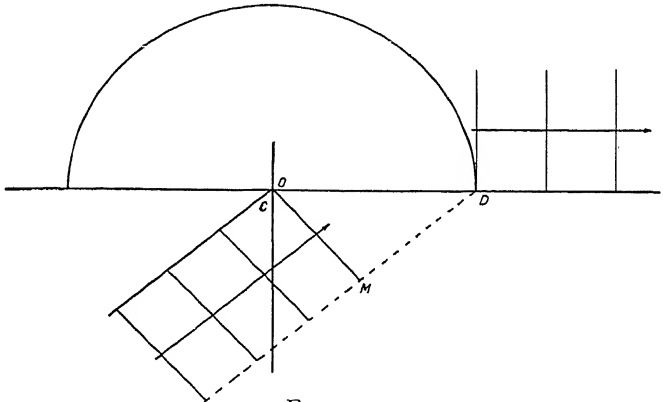


FIG. 40.

the case of air-water, we have this critical condition when $OD = 1.33MD$, or more generally when $\frac{OD}{MD} = \frac{v_{\text{air}}}{v_{\text{water}}} = n$.

But from Fig. 40, $\frac{1}{\sin C} = \frac{OD}{MD}$, where C is the critical angle of incidence, since $MOD = C$.

Therefore, we have the relation which the student should recognize,

$$\sin C = \frac{1}{\text{index of refraction}} \tag{3.01}$$

The question of total reflection may be approached somewhat differently. Suppose *spherical* waves spread out from a source P, Fig. 41, strike the plane surface AB separating two media, the

velocity of the wave-motion in the upper being greater than in the lower. (Fig. 41 has been drawn on the supposition that the ratio of velocities = 1.5.) As a wave-front advances from *P* it will strike the surface first at the point *O* causing a secondary wavelet to originate at this point. As time goes on, other wavelets will originate successively at particles 1, 2, 3, 4, 5, etc., so that, by the time the original wave-front has just reached the point 10,

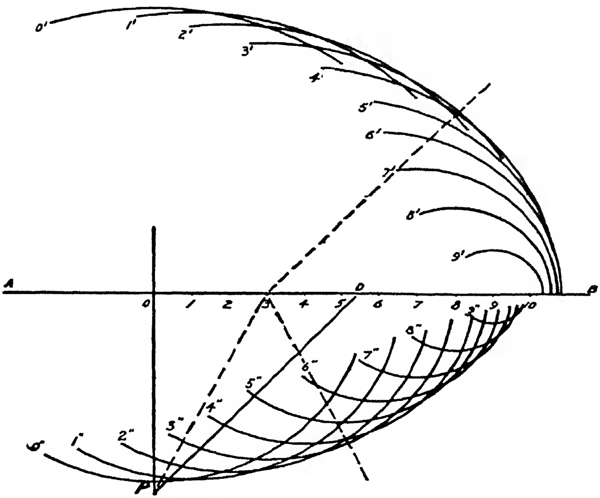


FIG. 41.

the reflected and the refracted wave-fronts may be found by drawing the surfaces tangent to all the wavelets in their positions *at that instant.* The student will readily see that the radius of the refracted wavelet originating in point 4, for example, $= (P10 - P4) \times 1.5$, while that of the corresponding reflected wavelet $= P10 - P4$. In this way wavelets for all the particles from 0 to 10 have been drawn.

A glance at the diagram will show that while there is an enveloping surface for all the reflected wavelets, and for the refracted wavelets from 0, 1, 2, 3, 4, and 5, *there is no real envelope for wavelets from 6, 7, 8 and 9.* This means that for all particles on

the surface beyond a point D somewhere between points 5 and 6, *no refracted wave-front is possible* because the individual wavelets fall within each other. It follows, then, that OPD is the critical angle for these two media. For purposes of reproduction the points 1, 2, 3, 4, etc., have been taken fairly far apart and an exact location of D is not possible. In the figure, the angle $OP5$ is $39° 44'$, $OP6$ is $44° 46'$, while the critical angle as calculated from formula (3.01) is 42°. It will be noted that not only have we

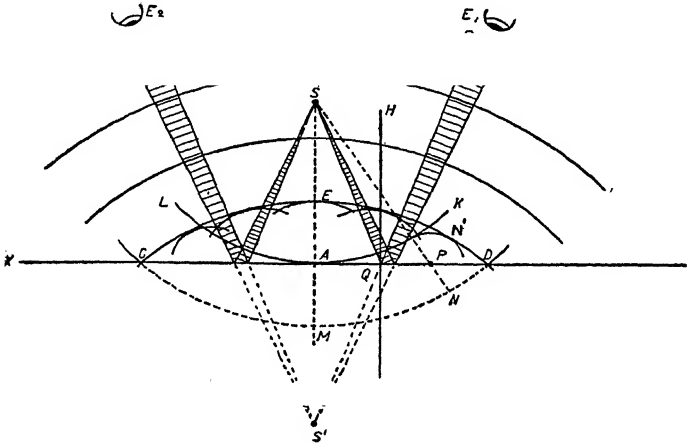


FIG. 42.

obtained a physical reason for total reflection but we have been able to make a fair estimate of the critical angle without the use of the sine law of refraction.

29. Spherical Waves Reflected at a Plane Surface.—Although in the above discussion of total reflection some consideration of reflection of spherical waves was necessary, this case is so important that we shall now consider it more carefully. Suppose, then, that spherical waves coming from a source S, Fig. 42, strike the plane reflecting surface XY. At a certain instant the incident wave-front will have reached the position LAK, and a wavelet will be on the point of spreading out from A. As the wave-front

advances, successive particles in the surface XY will become the source of other wavelets. In this figure we have represented the position of such wavelets at the particular instant that the original wave-front would have reached CMD had there been no reflecting surface. The wavelet originating at P, for example, will have a radius $= PN$. Applying Huygens' Principle, we see the formation of a reflected wave-front CED which will continue to propagate itself in the normal way. The reflected wave-front evidently leaves the reflecting surface *as if* it came from a centre S', the *virtual image* of S. Since $AE = AM$, $PN' = PN$, etc., the curvature of CED is equal to that of CMD, and, without any further formal proof, we see that this virtual image is as far behind the reflecting surface as the original source is in front — the second elementary law of reflection.

If the eye is in a position such as E_1, it "sees" the image by means of the disturbance confined within the shaded cone, apparently coming from S', in reality from the original source along the shaded path. In another position such as E_2, the same image is visible, but by means of the different cone shown in the figure.

If we deal with a ray, such as SQ, it is reflected along the path QB, and it is a simple matter to prove geometrically that the angle of incidence SQH = the angle of reflection HQB, where HQ is normal to the surface at Q.

30. Spherical Waves Refracted at a Plane Surface. — Here again we consider more in detail the case of the refracted wave-front arising from wavelets originating in particles 0, 1, 2, 3, 4, 5, Fig. 41. Suppose spherical waves spreading out from the source S, Fig. 43, pass through the plane surface XY into a less dense medium, where the velocity v_1 is greater than v_2, that in the lower. In the figure the positions of the refracted wavelets have been drawn at the instant an incident wave-front LMK would have reached the position ANB, had there been no change of medium. The wavelet from the point M, therefore, has a radius

$$MN' = MN \cdot \frac{v_1}{}.$$

Again Huygens' Principle shows us the existence of a refracted wave-front $AN'B$ spreading out from the surface as if it came from some centre S', the virtual image of S. In the case considered, S' is obviously nearer the surface than the original source. How much nearer, we shall find out in section 32. As in the

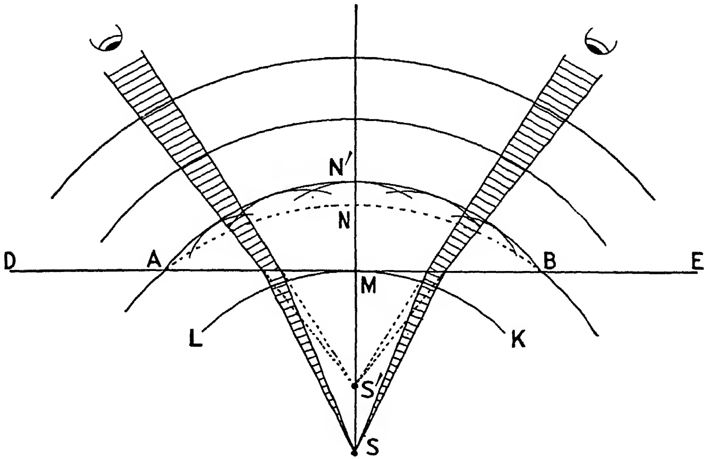


FIG. 43.

previous case of reflection, the eye sees the image in different positions by means of little cones of rays, two of which are shown in Fig. 43.

31. Before considering this case quantitatively, it is necessary to digress in order to establish an expression for measuring the curvature of a circle or sphere. In Fig. 44, DCE is a portion of circle or sphere of radius r; AOC is a diameter, and DBE a chord at right angles to this diameter.

By elementary geometry,

$$BD \cdot BE = AB \cdot BC$$

or

$$\begin{aligned} y^2 &= (2r - x)x \\ &= 2rx - x^2 \end{aligned} \tag{3.02}$$

where $y = BD = BE$, and $x = BC$.

Now, if x is a small quantity compared with y or r, with little error we may neglect its square, and write

$$y^2 = 2rx,$$

or
$$\frac{1}{r} = \frac{2x}{y^2} \tag{3.03}$$

Relation (3.03), of which we shall make considerable use, tells us that the curvature (the reciprocal of the radius) of a circle or a sphere is proportional to the distance x, sometimes called the sagitta, *provided x is sufficiently small.*

Many students have a feeling of dissatisfaction in thus dropping quantities, as is so frequently done in physics. Their difficulty lies in their failure to appreciate that physics is a science based on measurement. Now there is a limit of accuracy beyond which it is not possible to go in any measurement. If, for example, a student is measuring the distance of an image from a mirror, he may record 20 cm. That does not mean 20.00000 · · ·, for there are many factors which make it impossible for him to go beyond a certain degree of accuracy. He may be using a ruler graduated in millimeters, fractions of which he must estimate; the ruler may be unevenly scaled, the image may not be sharply focussed, and, in general, he finds his independent measurements varying somewhat. If asked to consider his measurements and use his judgment, he may finally say that his measurement cannot be relied on to more than 1 mm. His length 20 cm., therefore, is considered correct to about 1 part in 200, or to $\frac{1}{2}$ per cent, and might be written 20.0 ± 0.1 cm.

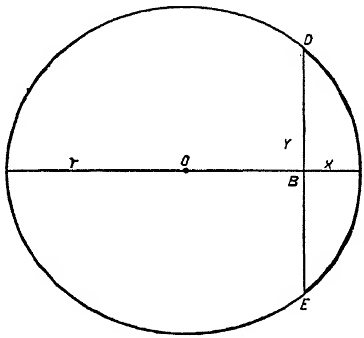


FIG. 44.

Now suppose this number has to be used in making calcula-

tions from a formula such as (3.02), which could be simplified by dropping x^2 to the simpler relation (3.03). If the dropping of x^2 means introducing an error considerably less than $\frac{1}{2}$ per cent, the student would be justified — indeed, he would only be using his common sense — in making use of the simpler expression.

The actual error in dropping x^2 may readily be found for special cases. For example, if $r = 60$ cm., $x = 5$ mm.,

$$\begin{aligned} y^2 &= 59.75, \text{ from relation (3.02)}, \\ &= 60.00, \text{ from relation (3.03)}. \end{aligned}$$

It will be seen that the magnitude of the error depends on the value of x compared with the other quantities. Whenever we use relation (3.03), we are assuming that this error may be neglected and it is important to realize that any relations which are derived from it, such as (3.05), hold only subject to this qualification.

32. We return now to the case of refraction illustrated in Fig. 43. Applying relation (3.03), we may write

$$\frac{1}{SN} = \frac{2MN}{MB^2} \quad (a)$$

and

$$\frac{1}{S'N'} = \frac{2MN'}{MB^2} \quad (b)$$

provided MN and MN' are small compared with SN and $S'N'$.

If an eye looking vertically downwards views an object below the surface of water, either through a microscope or unaided, the effective portion of the emergent wave-front will be extremely small. For example, in Fig. 43, if the source S is 25 cm. below the surface DE, and we take the aperture of the eye = 3 mm., the distance AB is less than 3 mm. In such a case the above condition will be fulfilled. It is well to realize that Fig. 43 is grossly distorted, MN being tremendously magnified.

Combining (a) and (b), we have

$$\frac{SN}{S'N'} = \frac{MN'}{MN}$$

$$= \frac{v_1}{v_2} = n,$$

or
$$\frac{SM + MN}{S'M + MN'} = n,$$

or
$$\frac{SM\left(1 + \frac{MN}{SM}\right)}{S'M\left(1 + \frac{MN'}{S'M}\right)} = n,$$

or
$$n = \frac{SM}{S'M}, \text{ with small error,}$$

or
$$= \frac{\text{object distance}}{\text{image distance}}. \tag{3.04}$$

In the case of water ($n = 1.33$), an object when viewed vertically appears to be about three-quarters of its actual depth below the surface.

33. We pass now to a consideration of curved reflecting and refracting surfaces. In all cases we shall assume that the surfaces are portions of spheres, and that the portion actually used is such a small fraction of the whole that we may use the approximate relation (3.03) as a measure of the curvature. The relations we deduce will then be applicable only when this restriction holds good. It is well to realize, however, that in our diagrams, for the sake of clearness, the lengths corresponding to x of Fig. 44 are all grossly exaggerated.

Plane Waves at a Spherical Surface. — Let ABC, Fig. 45, be a plane wave-front incident on the concave spherical mirror AOC, whose centre of curvature is R and radius of curvature $= r = RO$. Proceeding in the same way as has already been done more than once, we have drawn the position of wavelets at the instant the wave-front would have reached the position MON. This means that the radius of the wavelets originating

in A and in $C = BO$. Again we see the formation of a reflected wave-front $A'OC'$ which, since we are dealing only with a small portion of a curved surface, may be considered as spherical, making for a centre F, the *principal focus* or *focal point*.

The position of F evidently depends on the curvature of the mirror; for the larger BO, the greater the curvature of the reflected wave-front, that is, the shorter the focal length, OF.

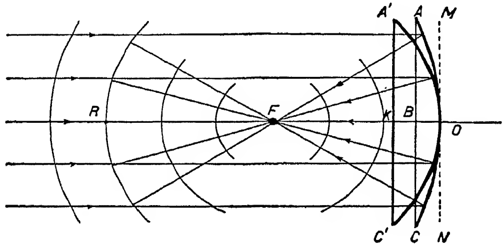


FIG. 45.

Join A' and C', the outer edges of the reflected wave-front. $A'C'$ will then intersect RO at right angles, and $KB = AA'$.

Since the incident disturbance advances from B to O, while the reflected goes from A to A',

$$AA' = BO,*$$

or

$$KB = BO.$$

$$KO = 2BO$$

$$\frac{2KO}{y^2} = 2 \quad \frac{2BO}{y^2},$$

multiplying each side by $\frac{2}{y^2}$, where $y =$

$A'K = AB = \frac{1}{2}$ the aperture of the mirror.

$$\therefore \quad \frac{1}{OF} = \frac{2}{OR}, \text{ by relation (3.03)}$$

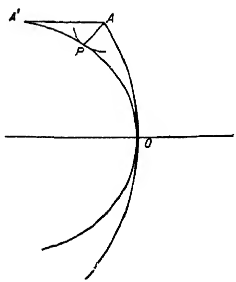


FIG. 46.

* It may be objected that it is not the distance AA' but AP in Fig. 46 which is equal to BO. This is quite true, but for sufficiently small apertures the angle $A'AP'$ will be so small that its cosine = 1, and with small error $AA' = AP$. Fig. 46 is, of course, grossly exaggerated.

or

$$f = \frac{r}{2}, \tag{3.05}$$

that is, the focal length is equal to one-half the radius of curvature — subject always to the restriction of small aperture. If the aperture is large, the focal length of the central portion differs somewhat from the outer portions, and plane waves from a distant

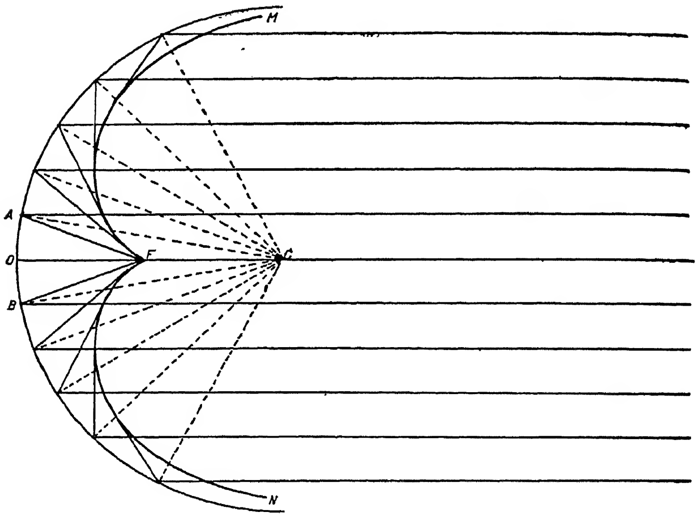


FIG. 47.

point give rise not to a point image but to the familiar caustic curve *MFN* illustrated in Fig. 47.

It is left to the student to prove relation (3.05) for a convex mirror.

34. Spherical Waves Reflected from a Spherical Surface. — Plane waves imply distant sources, and in actual practice we more frequently have to do with sources which are not distant. Suppose, then, that spherical waves spread out from a point source *P*, Fig. 48, fall on a concave mirror *AOC*, whose centre of curvature

is at R. Let ABC be an incident wave-front, striking the mirror first at A and C. Without going into details, the student should now be able to see that $A'OC'$ will be a reflected wave-front, with $AA' = BO$. This reflected wave-front will advance towards a centre Q, the real image of P, from which the disturbance will continue to spread out in ever-widening areas.

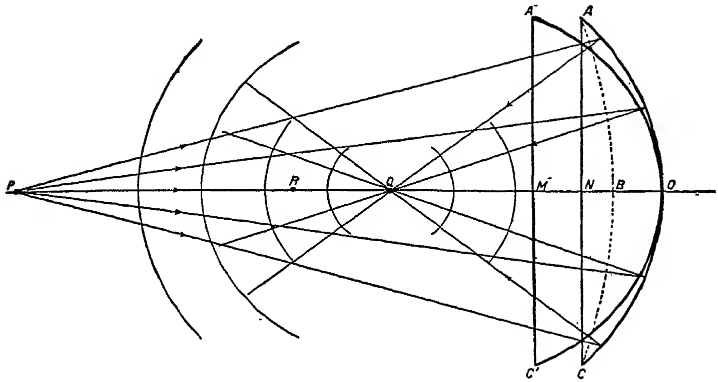


FIG. 48.

Let $p = PO$, the distance of object P from the mirror;
$q = OQ$, the corresponding distance of image Q.

We wish to find a relation connecting p, q, and r.

Since $AA' = BO$,
$\therefore\ MN = BO$,

M and N being points of intersection of $A'C'$ and AC with the axis PQO

$$\therefore\quad MO - NO = NO - NB.$$

This step is taken to introduce only those distances which are proportional to the curvatures $\frac{1}{p}$, $\frac{1}{q}$, and $\frac{1}{r}$.

$$\therefore \quad MO + NB = 2NO$$

or $$\frac{2MO}{y^2} + \frac{2NB}{y^2} = \frac{2 \cdot 2NO}{y^2} \quad \text{where } 2y = \text{aperture.}$$

$$\frac{1}{q} + \frac{1}{p} = \frac{2}{r}$$

$$= \frac{1}{f} \qquad (3.06)$$

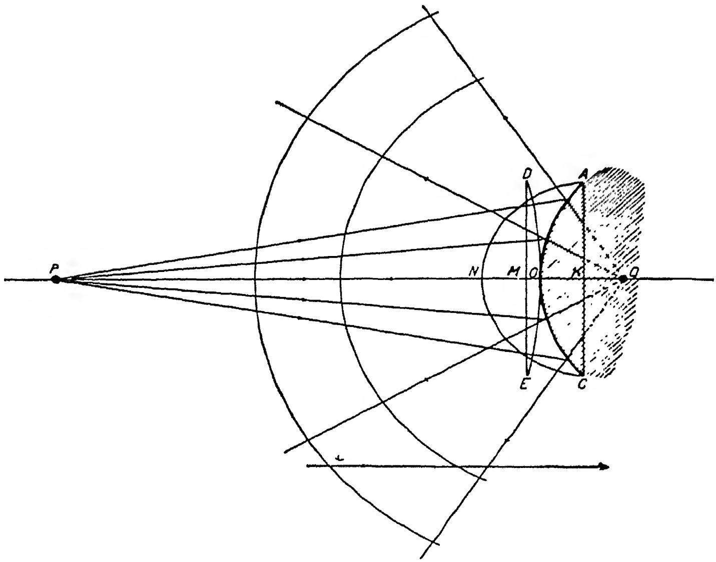


FIG. 49.

35. Spherical Waves Reflected from a Convex Mirror. — As in the previous case, AOC, Fig. 49, represents the mirror, DOE an incident wave-front from the source P, ANC the reflected wave-front, while DE and AC intersect the axis POQ at M and K. In this case the central portion of the incident wave-front strikes the mirror first, and the reflected wave-front comes off *as if* it came from the image Q, which is therefore virtual.

Since

$$AD = NO,$$
$$MK = NO$$
$$MO + OK = NK - OK$$
$$\frac{2MO}{y^2} - \frac{2NK}{y^2} = -2 \cdot \frac{2OK}{y^2}$$
$$\frac{1}{p} - \frac{1}{q} = -\frac{1}{f}. \tag{3.07}$$

It will be noticed that the signs in relation (3.07) differ from those of (3.06). Since these relations are frequently used, it is desirable, if possible, to have a single relation applicable to both kinds of mirrors. This can be done provided some sign convention is adopted. The student will find that all writers do not adopt either the same final relation or the same convention. In this book the convention which is used is the same for all cases of reflection and refraction at spherical surfaces, and, moreover, it leads to the description of a converging lens as positive, a diverging lens as negative, in accordance with the agreement of the international convention of opticians. According to it, distances measured from the mirror or lens or surface in the direction in which the light is travelling are positive; distances measured in the opposite direction, negative. In other words, if, in a diagram, the light travels from left to right, and the centre of the mirror or lens or curved surface is placed at an imaginary origin of co-ordinates, the axis of the optical system coinciding with the x-axis, distances on the right of the origin are positive, on the left, negative.

Adopting this convention, we then write

$$\frac{2MO}{y^2} \qquad \frac{1}{p}, \quad \frac{2NK}{y^2} = +\frac{1}{q}, \text{ and } \frac{2OK}{y^2} \qquad \frac{1}{r}.$$

Instead of relation (3.07) we then have

$$\frac{1}{-p} - \frac{1}{+q} = -\frac{1}{+f}$$

or $\frac{1}{q} + \frac{1}{p} = \frac{1}{f}$, a relation exactly the same as that (3.06) deduced for the concave mirror.

In using relation (3.06) in a numerical problem, the student must never forget to substitute values subject to the sign convention. A simple example may not be out of place.

A small object is placed 50 cm. from a convex mirror whose radius of curvature is 20 cm. Find where the image is.

We are here given $p = -50$, $f = +10$,

since $$\frac{1}{q} + \frac{1}{p} = \frac{1}{f},$$

we have $$\frac{1}{q} + \frac{1}{-50} = \frac{1}{+10}$$

or $$q = +8.3 \text{ cm.}$$

The plus sign means that the image is on the side of the mirror opposite to that from which the light comes, that is, is a virtual image 8.3 cm. behind the mirror.

36. Refraction at a Spherical Surface. — The same general method may be used to deduce relations connecting object and image distances when refraction takes place at a spherical surface separating two media. Two such cases will be discussed.

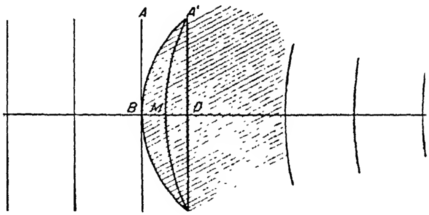


FIG. 50.

(*a*) *Plane waves incident on a convex surface.* In Fig. 50, $A'BC'$ represents a convex surface, radius of curvature $= r$, separating the less dense medium on the left from the more dense (shaded) on the right, while ABC is an incident plane wave-front. Since the centre of the incident wave-front strikes the surface first at B, a wavelet will begin to spread out from this point,

while points between B and A' will be successively disturbed. To obtain the refracted wave-front, for example, at the instant the outer edges of the plane wave have reached the refracting surface at A' and C', all that is necessary is to draw the position of the wavelet which originated at B. If BM is the radius of this wavelet at the instant under consideration, then $A'MC'$ is the new refracted wave-front, making for some centre Q — the real image.

Because of the slower velocity in the more dense medium, its radius BM will be less than the distance AA' in the ratio of the two velocities, and we may write

$$AA' = n \cdot BM, \quad (a)$$

where n = the index of refraction for these two media.

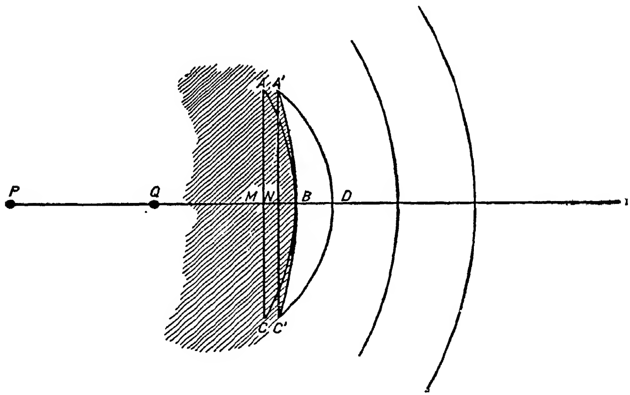


FIG. 51.

To find the distance of the image from the surface, all that is now necessary is to obtain from this equation a relation involving the distances proportional to the various curvatures. Thus we now write instead of (a),

$$BD = n \cdot BD - n \cdot MD$$

or

$$n \cdot MD = (n - 1) \cdot BD$$

or
$$\frac{2n \cdot MD}{y^2} = (n - 1) \cdot \frac{2BD}{y^2}$$

from which
$$\frac{n}{f} = \frac{n-1}{r}$$

or
$$\frac{1}{f} = \frac{n-1}{n} \cdot \frac{1}{r} \tag{3.07}$$

(*b*) *Spherical waves emerge from a more dense medium through a spherical surface into a less dense medium.* In Fig. 51, ABC is the incident spherical wave-front, $A'BC'$ the spherical refracting surface of radius of curvature r, while $A'DC'$ will be the emergent refracting wave-front if $n \cdot AA' = BD$.

We have, therefore,
$$n \cdot MN = BD$$

or
$$n \cdot MB - n \cdot NB = ND - NB$$

or
$$n \cdot MB - ND = (n - 1)NB$$

or
$$\frac{n}{-p} - \frac{1}{-q} = \frac{n-1}{-r}$$

or
$$\frac{n}{p} - \frac{1}{q} = \frac{n-1}{r}, \tag{3.08}$$

where p and q stand for object and image distance.

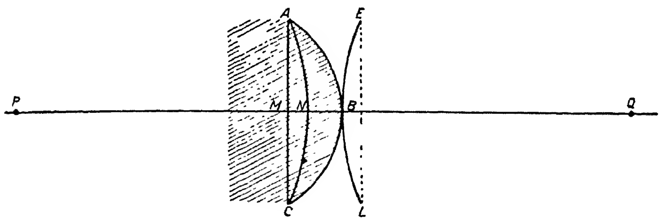


FIG. 52.

If the curvature of the surface is sufficiently large, and the object at some distance from the surface, the emergent wave-front EBL, Fig. 52, may have the opposite curvature to that of the incident ANC. It is left to the student to prove, that, subject always to the sign convention, the same relation holds.

If, using the notation given at the end of section 27, we designate the index of the medium in which the object lies by n_1 and the index for the image medium by n_2, relation (3.08) may be put into the more convenient form

$$\frac{n_2}{q} - \frac{n_1}{p} = \frac{n_2 - n_1}{r}. \tag{3.08a}$$

It is left as an exercise for the student to derive this relation by applying exactly the same method as has been used in the derivation of (3.08).

Relation (3.08a) can be made still more symmetrical by following the method, frequently used in geometrical optics, of designating an object distance by s_1 and an image distance by s_2. In that case, (3.08a) becomes

$$\frac{n_2}{s_2} - \frac{n_1}{s_1} = \frac{n_2 - n_1}{r},$$

a relation which is sometimes designated as the paraxial equation for refraction at a single spherical surface.

In this text, however, we shall continue to use the symbols p and q.

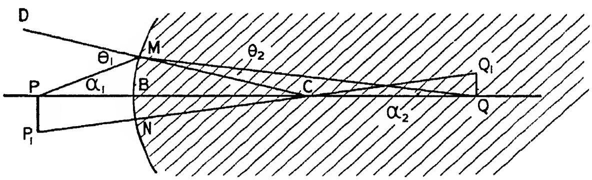


FIG. 52A.

36A. The Sine Condition. — Although it is not one of the aims of this text to discuss details of geometrical optics, a relation of great importance in that branch of light is derived in this section, largely because some use of it is made later in connection with the resolving power of a microscope.

In Fig. 52A, let PP_1 represent a small object in a medium of index n_1 and QQ_1 the corresponding image formed in the medium

of index n_2 after refraction at the curved surface MBN. The object distance is then $PB = p$; the image distance is $BQ = q$; and, if C is the centre of curvature of the refracting surface, $BC = r$, the radius of curvature.

Since Q is the image of P, then any paraxial ray incident at M must travel the path MQ in the image medium. Hence, since CMD is perpendicular to the surface at M,

$$\angle PMD = \text{angle of incidence} = \theta_1,$$

and

$$\angle QMC = \text{angle of refraction} = \theta_2.$$

As shown in section 27, it follows that

$$n_1 \sin \theta_1 = n_2 \sin \theta_2. \quad (a)$$

Moreover, if Q_1 is the image of P_1, then the ray from P_1 to Q_1 through C, the centre of curvature, is a straight line, being normal to the refracting surface at N.

Since the triangles PCP_1 and QCQ_1 are similar,

$$\frac{QQ_1}{PP_1} = \frac{QC}{PC}$$

or

$$\frac{y_2}{y_1} = \frac{q - r}{p + r} \quad (b)$$

where y_1 and y_2 are symbols for the lateral extensions of object and image.

From triangle PMC,

$$\frac{\sin MPC}{r} = \frac{\sin PMC}{p + r},$$

or, putting $\angle MPC = \alpha_1$, the *slope angle* of the ray PM in object space,

$$\frac{\sin \alpha_1}{r} = \frac{\sin \theta_1}{p + r},$$

or

$$p + r = \frac{r \sin \theta_1}{\sin \alpha_1}. \quad (c)$$

Similarly, from triangle QMC,

$$q - r = \frac{r \sin \theta_2}{\sin \alpha_2} \qquad (d)$$

where $\angle PQM = \alpha_2$, the slope angle of the ray MQ in the image space.

Combining (c) and (d), we have

$$\frac{- r}{p + r} \quad \frac{\sin \theta_2}{\sin \alpha_2} \frac{\sin \alpha_1}{\sin \theta_1}$$

$$\frac{n_1 \sin \alpha_1}{n_2 \sin \alpha_2}; \text{ from } (a). \qquad (e)$$

Finally, combining (b) and (e), we have

$$\frac{y_2}{y_1} = \frac{n_1 \sin \alpha_1}{n_2 \sin \alpha_2},$$

or
$$n_1 y_1 \sin \alpha_1 = n_2 y_2 \sin \alpha_2.$$

This is the sine condition.

This proof, it will be noted, applies only to a single refracting surface and hence, because the restriction of small aperture limits the position of M to points near the axis, the angles α_1 and α_2 are small. It is important to note, however, that the sine condition may be applied, even when these angles are not small, to any system of refracting surfaces which has been corrected for defects so that a point image is formed at Q for *all* rays which leave P and pass through the system. For example, it may be applied to the objective lense of a good microscope. (Section 56A.)

For a single refracting surface of small aperture, the sine condition reduces to

$$n_1 y_1 \alpha_1 = n_2 y_2 \alpha_2.$$

This is a statement of Lagrange's Theorem, according to which the product $ny\alpha$ is constant for any number of refractions, if attention is confined to paraxial rays.

Examples of the use of the sine condition will be found on pages 141, 143 and 286.

37. In any problems dealing with refraction at spherical surfaces, it is better to work not from a formula, but from fundamental considerations. When this is done, the sign convention may or may not be used. If it is used, it is important to realize that two steps are necessary: (1) to establish the general relation, (2) to substitute numerical values, — both subject to the convention. The following example will make the point clear.

Waves from a point source in air fall on a convex surface separating air from glass ($n = 1 \cdot 5$). If the source is 100 cm. from the surface, whose radius of curvature is 20 cm., find the position of the image.

Let ABC, Fig. 53, be the incident wave-front,
DBE the convex surface,
DNE the emergent wave-front, making for the image Q, q cm. from the surface.

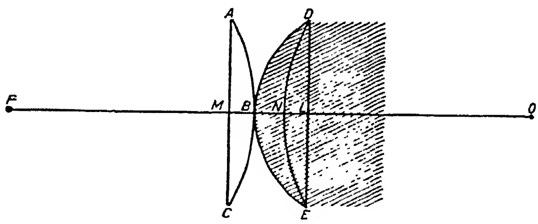


FIG. 53.

Since the central portion of the disturbance travels in glass from B to N, while the outer portions go in air from A to D, or C to E,

$$AD = 1.5BN$$
$$\therefore \quad ML = 1.5BN$$
$$\therefore \quad MB + BL = 1.5BL - 1.5NL$$
$$\therefore \quad 1.5NL + MB = (1.5 - 1)BL.$$

Without the sign convention we may write at once, since the student should now be able to multiply each side by $\frac{2}{y^2}$ mentally,

$$\frac{1.5}{} + \frac{1}{100} = \frac{0.5}{20}$$

from which $q = 100$ cm.

With the sign convention, we must first establish the general formula, thus

$$\frac{1.5}{+q} + \frac{1}{-p} = \frac{0.5}{+r} \quad (b)$$

We now use this formula (b) by substituting for $p = -100$, $r = +20$, and we then have

$$\frac{1.5}{+q} + \frac{1}{100} = \frac{0.5}{+20},$$

from which $q = +100$ cm.

The numerical answer in each case is, of course, the same. In the second case the plus sign tells us that the image is on the positive side of the surface.

It may seem to the student that the two steps taken in the second solution annul each other, and in one sense this is quite true. On the other hand the final answer must always bear a plus or a minus sign, and this enables us to say at once on which side of the surface the image is to be found.

38. Lenses. — In dealing with the passage of waves through lenses, we shall confine our attention to problems dealing with those whose surfaces are either plane or spherical, and whose thickness can be neglected in comparison with object and image distances. (See, however, section 47.) Only paraxial rays will be considered and, unless otherwise stated, we shall assume that the medium on both sides of the lens is always the same.

Although a variety of shapes is possible, it is well to note that all thin lenses may be divided into two groups, according as their thickness at the centre is greater or less than at the edges.

Plane Waves Incident on a Lens Thicker at the Centre. — (a) *Medium of lens denser than that of its surroundings.* In Fig. 54,

HBK and HEK represent the two spherical surfaces of such a lens, radii of curvature $= r_1$ and r_2; ABC is an incident wave-front; DEL the emergent wave-front making for a centre F, the *principal focus* or *focal point*. To find the value of f, the focal length, we follow the same method as in the case of mirrors and refracting surfaces. Thus, since the central portion of the disturbance travels in the lens from B to E, while the edges go from A to D (or C to L) in the surrounding medium, we have

$$AH + HD = n \cdot BE$$

$$\therefore \quad BM + MN = n \cdot BM + n \cdot ME$$

$$\therefore \quad BM + ME + EN = n \cdot BM + n \cdot ME$$

$$\therefore \quad EN = (n - 1)(BM + ME)$$

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r_1} + \frac{1}{-r_2}\right)$$

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \qquad (3.09)$$

This is an important fundamental thin lens relation.

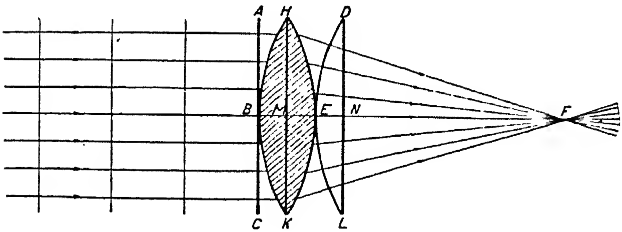


FIG. 54.

(*b*) *Lens medium less dense than that of the surroundings.* In this case, since the velocity in the lens medium is greater than in the surroundings, the central portion of the disturbance will get ahead of the edges, and a plane wave-front such as ABC, Fig. 55, will give rise to an emergent wave-front HNK, which apparently comes from the virtual principal focus F. This is

exactly what happens if a parallel beam of light is allowed to fall on an air lens of this shape (bounded by two thin glass walls) immersed in a tank of water.

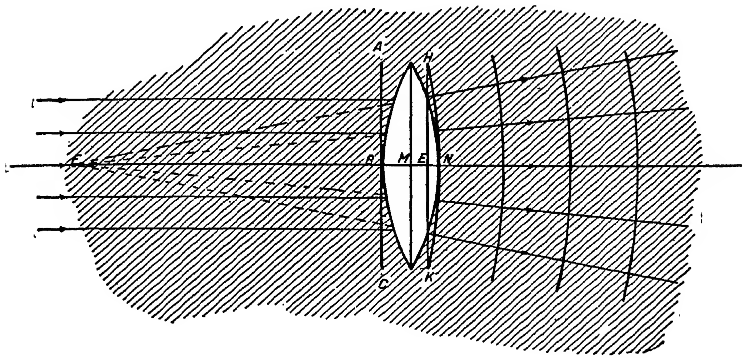


FIG. 55.

It is left to the student to prove for himself that relation (3.09) holds for this case, as well as when a lens thinner at the centre is used, provided n is always considered the index of the lens material with respect to the surrounding medium.

39. Spherical Waves Incident on a Thin Lens. — Fig. 56 represents the case of a more dense double-convex lens on which a wave-front ANB is incident. EHF is the emergent wave-front

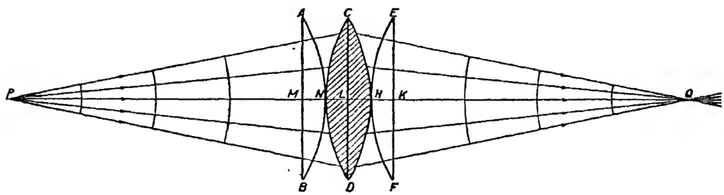


FIG. 56.

making for the real image Q, while CND and CHD are the lens surfaces with radii of curvature equal to r_1 and r_2 respectively.

We have then

$$AC + CE = n \cdot NH$$

$$\therefore \quad ML + LK = n \cdot NH$$

$$\therefore \quad MN + NL + LH + HK = n \cdot NL + n \cdot LH$$

$$\therefore \quad MN + HK = (n - 1)\ (NL + LH)$$

$$\therefore \quad \frac{1}{-p} + \frac{1}{+q} = (n - 1)\left(\frac{1}{r_1} + \frac{1}{-r_2}\right)$$

or

$$\frac{1}{q} - \frac{1}{p} = (n - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

$$= \frac{1}{f}, \qquad (3.10)$$

where p and q are object and image distances.

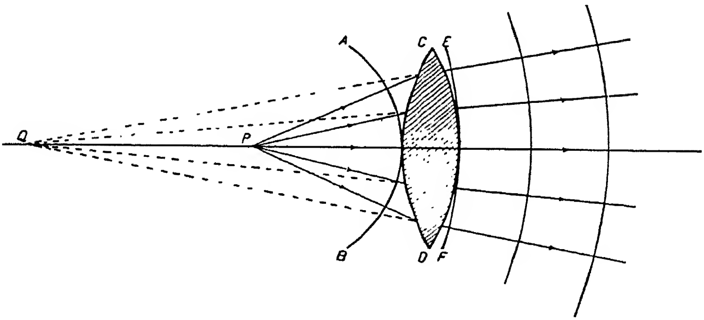


Fig. 57.

Relation (3.10) will be found to apply to any case of spherical waves, falling on any kind of lens. For example, in Fig. 57 we have given the diagram when the source P is so near a double convex lens that the curvature of the emergent wave-front is lessened but not reversed (as in Fig. 56), and the image is, in consequence, virtual.

Again, Fig. 58 is the diagram for a dense lens thinner at the centre, in which case the curvature of the incident wave-front

ABC is increased so that the emergent wave-front DEF comes off as if from the virtual image O.

40. Spherical Aberration. — It has already been emphasized that the relations we have derived apply only to cases where the aperture is small and paraxial rays are used. Relation (3.09), for example, gives the focal length of a thin lens only for a

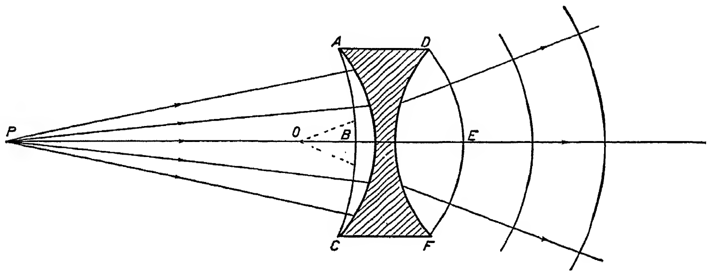


FIG. 58.

small portion near the axis. If rays, parallel to the axis but incident at a region distant h cm. from the axis, are considered, the focal length for this region differs from that given by (3.09) by an amount which is greater the greater the value of h. In other words, a wide beam of parallel rays cannot be brought to a single sharp focal point by a spherical lens of wide aperture. This variation in the focal length of a lens when we pass from the central portion of the periphery is called *spherical aberration*. It is possible to correct a lens for this defect by grinding suitable aspherical surfaces, but since spherical surfaces are the only kind which can be worked commercially in quantity, single lenses invariably have this defect.

Spherical aberration, however, may be reduced to a minimum by a suitable choice of the curvatures of the faces of a lens. The point is illustrated by the following examples in which the numbers are taken from a table in *The Principles of Optics* by Hardy and Perrin.*

* McGraw-Hill Book Company, Inc., 1932.

In each example the aberration, which is measured by the difference between the focal length of the axial region and that of a small region 1 cm. from the axis, is for a lens of focal length 10 cm.

(*a*) For a plano-convex lens, with plane face next the object and radius of curvature of the other face 5.2 mm., the aberration is 4.4 mm. When the lens is reversed, so that the curved face is next the object, the aberration is 1.1 mm.

(*b*) When the shape of the lens is double-convex, with each face having a radius of curvature of 10 cm., the aberration is 1.5 mm.

When using a single plano-convex lens the advantage of turning the curved face towards the object is obvious. In general, it may be stated that the spherical aberration for a given lens is least so placed that the deviation of a ray at each face is as nearly equal as possible.

By suitably choosing combinations of lenses, spherical aberration may also be lessened. When discussing eyepieces in section 55, reference will again be made to this point.

41. In actual practice we almost always have to do with lenses made of a material more dense than the surroundings. When this is so, it will readily be seen that lenses thicker at the centre than at the edges are always converging, while those thinner at the centre are always diverging.

It will be noted too that according to the sign convention used in this book, the focal length of a converging lens is positive, that of a diverging lens, negative.

42. We conclude this chapter with a summary of the important general relations, together with one or two examples of their use in solving problems.

(*a*) $\frac{1}{q} + \frac{1}{p} = \frac{1}{f}$, applicable to all (spherical) mirrors. (3.06)

(*b*) $\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$, applicable to all (spherical) lenses. (3.10)

(*c*) $\frac{1}{f} = (n - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$ applicable to all (spherical) lenses. (3.09)

All relations subject to the sign convention.

Problem 1. *An object is placed 10 cm. in front of a concave mirror whose radius of curvature is 30 cm. Find the position and nature of the image.*

Here we are given $p = -10$ cm. and $f = \frac{r}{2} = -15$ cm.

∴ substituting in the formula for spherical mirrors

$$\frac{1}{q} + \frac{1}{p} = \frac{1}{f},$$

we obtain $q = -30$ cm.

The image, therefore, is virtual and is situated 30 cm. from the mirror on the same side as the object.

Problem 2. *The radii of curvature of the faces of a biconvex lens are 20 cm. and 30 cm. Its focal length is 24 cm. What is the index of refraction of the glass?*

Here we are given $f = +24$ cm.; $r_1 = +20$ cm.; $r_2 = -30$ cm.

∴ substituting in the general formula

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right),$$

$$\frac{1}{24} = (n - 1)\left(\frac{1}{20} - \frac{1}{-30}\right)$$

from which

1.5.

Problem 3. *A convex glass lens has a focal length of 40 cm. in air. What is the focal length when immersed in water? Index for air – glass = 1.54; for air – water = 1.33.*

For glass lens in air, we are given $n = 1.54$, $f = +40$,

∴ substituting in the general formula, we have

$$\frac{1}{40} = (1.54 - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \quad (a)$$

For the same glass lens in water, we are given $n = \frac{1.54}{1.33}$, and therefore can write

$$\frac{1}{f} = \left(\frac{1.54}{1.33} - 1\right)\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \quad (b)$$

Combining equations (a) and (b), we find the required $f = +137$ cm.

Problem 4. *An object is placed 100 cm. from a thin diverging lens of focal length 20 cm. Find where the image is.*

In this problem we must use the general formula

$$\frac{1}{q} - \frac{1}{p} = \frac{1}{f}.$$

We are given $p = -100$, $f = -20$, and hence by substitution readily find $q = -16.7$ cm.

We have therefore a virtual image 16.7 cm. from the lens on the negative side.

Problems

1. The water in a certain vessel is 12 in. deep. What is the apparent depth to an eye looking vertically down upon its surface?

2. What is the velocity of light in water? n for air-water $= \frac{4}{3}$.

3. An incandescent lamp is placed 6 ft. below the surface of a pond. Show why only a fractional part of the light can escape directly from the water.

4. If a beam of light has 5000 light waves to an inch in air, how many to the inch will there be after it has entered water?

5. Find the velocity of light in water, if the critical angle at the surface between water and air is 48° 30′.

6. When the index of refraction of water is 1.33 and that of carbon bisulphide is 1.67, what is the critical angle between water and carbon bisulphide?

7. A layer of sulphuric ether (index = 1.36) 2 cm. deep rests on water (index = 1.33) 3 cm. deep. What is the apparent distance of the bottom of the vessel below the surface?

8. A source emitting waves travels through a medium with a velocity v which is greater than the velocity V with which the wave-disturbance travels in this medium. (*a*) Apply Huygens' Principle to show that a conical wave-surface will be produced. (*b*) Find also the slope of the cone. (*c*) Give an illustration of the above in the case of sound or water waves.

9. Is there any limitation to the truth of the statement that the focal length of a concave mirror is equal to one-half its radius of curvature? Explain fully.

10. An air bubble imprisoned in a piece of glass appears to lie 0.5 cm. below the surface of the glass, which has a convex surface of 20 cm. radius. If the refractive index of the glass is 1.5, what is the real distance of the air bubble below the surface?

11. Prove, by method of change of curvature of wave-front, that, when two thin lenses of focal lengths f_1 and f_2 are placed in close contact, the resultant focal length is given by the relation

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}.$$

12. The curved surface of a plano-convex lens has a radius of curvature of 10 cm. What is its principal focal length when submerged in water? n for air-water $= \frac{4}{3}$; for air-glass $= \frac{3}{2}$.

13. The radii of curvature of the surfaces of a biconvex lens are 20 and 30 cm. Its principal focal length is 24 cm. What is the index of refraction of its glass?

14. Prove graphically that the position of the image of a distant object formed by a thin converging lens does not move when the lens is rotated about an axis through the centre, through a small angle.

15. A glass lens has a focal length of 200 cm. in water and is converging. What is the focal length of an air lens of the same shape and dimensions in water? Is it converging or diverging?

Refractive index of glass = 1.55.

Refractive index of water = 1.33.

16. A lens has two convex spherical faces, the radius of one is 15 cm., that of the other is 20 cm. The index of refraction of the glass is 1.5. What is the principal focal length of the lens?

17. A double convex lens with faces each of radius of curvature 20 cm. closes the end of a long tank filled with water. If a point source is placed outside the tank 100 cm. from the lens at a point on its axis, find where the image will be. Use the method of change of curvature of wave-front, not a formula.

Index of refraction: air-glass = 1.5; air-water = 1.3.

18. Light from a point source 30 cm. from a thin converging lens of focal length 20 cm. on emerging from the lens falls on the plane face of a large glass block. If the face of the block is 10 cm. from the lens find, by the method of change of curvature, to what point in the glass block the light is focussed. Index air-glass = 1.5.

19. Plane waves fall on a glass double convex lens (radius of curvature of each face = 20 cm., index = 1.5) which floats on clean mercury. Find, by applying the method of change of curvature of the wave-front, where the image is finally formed.

20. Find at what distance from a convex surface, of radius of curvature r, separating air from glass, a point source must be placed (in air) so that plane waves traverse the denser medium. Index = n.

21. Plane waves traversing a dense medium are incident on a spherical surface separating this medium from air. Find the position of the image. Radius of curvature of surface = r; index = n.

22. Spherical waves spreading out from a point source in air are incident on a convex surface separating air from a denser medium. If the source is p cm. from the surface, and a real image is formed q cm. from it prove that

$$\frac{n}{q} - \frac{1}{p} = \frac{n-1}{r},$$

where n is the index and r the radius of curvature of the surface.

23. A thick piece of glass (index = 1.5) has one face plane, the other spherical and curved outwards, with radius of curvature 20 cm. If a small object is placed outside the block at a point on the axis 10 cm. from the plane face, find where the image is, as viewed through the block. Thickness of block along axis = 5 cm.

24. A glass double convex lens, each of whose faces has a radius of curvature 20 cm., is surrounded by air on one side, by water on the other. If plane waves are incident on the air side, find (by the method of change of curvature of wave-front) to what point the waves are focussed.
Index air-glass = 1.6; air-water = 1.4.

25. (*a*) A thin double-convex glass lens has faces whose radii of curvature are 30 cm. and 20 cm. respectively. Find, *by the method of change of curvature of wave-front*, where a point object should be placed (on the side of the lens next the 30 cm. face) so that the wave-front just emerging from the lens coincides with the 20 cm. face. Index for air-glass = 1.5.
(*b*) Check your solution by finding (i) the value of the focal length of the lens, (ii) the position of the image (using p, q, and f).

26. A long tank is separated into two compartments by a thin double-concave air lens, bounded by very thin glass walls whose radii

of curvature are 20 cm. One compartment is filled with water, index 1.33, the other with a transparent liquid of index 1.21. If a parallel beam of light traversing the water strikes the lens, find by method of change of curvature of wave-front, where it is focussed.

27. A double-convex lens (index of refraction = n) whose faces have radii of curvature R_1 and R_2 floats on clean mercury. A small object is placed p cm. from the upper surface of the lens, this distance being such that the image formed by the rays which retrace their paths after reflection at the mercury surface coincides with the object. Prove, using method of change of curvature of wave-front (without sign convention) that

$$\frac{1}{p} = \frac{n-1}{R_1} + \frac{n}{R_2}.$$

28. A small object embedded in glass is viewed normally through the upper surface, which is convex upwards and has a radius of curvature 20 cm. The object is 2 cm. below the surface and the index of glass is 1.56.

Two students calculate the position of the image of the object, one treating the surface as if it were plane, the other allowing for the curvature. Find the per cent. difference in their values.

29. Parallel rays of light after traversing a tank of water are incident on a double-convex glass lens, through which the light emerges into air. Find by the method of change of curvature of wave-front, where the image is formed. Radii of curvature of lens faces are 30 cm. (next water) and 20 cm.; index air-glass = 1.56; air-water = 1.33.

30. Plane waves travelling in air fall on a convex surface, with radius of curvature 30 cm. separating air from glass. Find, by method of change of curvature, at what point the waves are brought to a focus. Index for glass = 1.55.

31. A small object is placed 100 cm. from a concave surface, of radius of curvature 20 cm., separating air from glass (index = 1.6). Find by the method of change of curvature of wave-front where the image is.

32. Plane waves fall normally on a double convex glass lens (index = 1.524), each of whose faces has a radius of curvature of 20 cm. The second face of the lens is silvered so that the waves are reflected. Find by the method of change of curvature to what point the waves are focussed.

33. Solve the preceding question, when the second face of the lens is plane, the other constants remaining the same.

34. Find the curvature of the curved face of a plano-convex lens, if its focal length is 20.0 cm. and the index of refraction of the lens material is 1.524.

35. Light from a small object in air falls on a convex surface separating air from glass ($n = 1.5$). If, when the object is 100 cm. from the surface, an image is formed at a point in the glass 100 cm. from the surface, find the linear magnification of the image. (Assume paraxial rays and use the sine condition.)

36. A small object is placed 30 cm. from a thin lens of focal length 20 cm. By the use of the sine condition, find the linear magnification of the image

CHAPTER IV

FURTHER STUDY OF LENSES

43. In Chapter III we have seen how, by a consideration of waves and Huygens' Principle, it is possible to show the formation of images by mirrors and lenses. While the method is applicable to waves in general, in this book we are primarily interested in light waves. Now, in light, especially in problems relating to the use of such combinations of lenses as are found in telescopes, microscopes, and other optical instruments, it is frequently more convenient

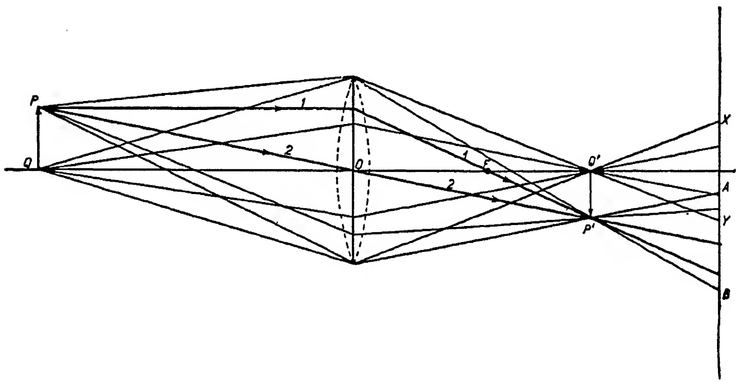


FIG. 59.

to deal with rays. In Figs. 54, 55, 56, 57, and 58, therefore, bundles of rays have been shown in accordance with the definition of a ray given on p. 49. The student is advised to examine these diagrams again carefully so that, in dealing with the formation of images by lenses, he will think rather of whole bundles of rays than of the principal rays by means of which such images may be located geometrically. For example, in Fig. 59, P' the image of the point source P has been located by finding the point of inter-

section of: (1) ray 1, which, incident parallel to the axis, emerges through the principal focus F; (2) ray 2, which goes through the centre of the lens unchanged in direction. It is important to realize, however, that a whole cone of rays leaves the source P, completely filling the aperture of the lens, and that after emergence, a corresponding cone goes to the image P'. Similarly for Q and for every other point on the object, there will be other complete cones falling on the lens, and eventually making for corresponding point images.

If a screen be placed at $P'Q'$ a sharp image will be seen, which will be *brighter* the wider the aperture, since a wider lens collects a larger bundle of rays. (Because of spherical aberration, the image may not be so *sharp*.) If there is no screen, the rays continue to spread out from each point on the image, such as Q' and P'. Such images will be seen by an eye placed anywhere in the region AB, for P', in the region XY for Q', and so on for other points. A clear understanding of where an eye must be placed in order to see an image is so important that a numerical example will be given.

44. *A point source is placed at a point on the axis 3 cm. from a convex lens of focal length 5 cm. and aperture 2 cm. Find through what distance an eye (considered as a point) may be moved in a direction at right angles to the axis, at a distance 2 cm. from the lens, and in all positions see the image.*

The *position* of the image may at once be calculated from the relation $\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$, for we are given $p = -3$, $f = +5$, and by substitution find $q = -7.5$ cm.

In Fig. 60, therefore, $PO = 3$, $QO = 7.5$, $MN = 2$, and the incident cone of rays leaves the lens as if coming from Q. An eye moved along the line XLY evidently can see the image Q only when between the points H and K.

Moreover, since triangles QOM and QHL are similar,

$$\frac{HL}{QL} = \frac{OM}{QO}$$

or $$\frac{HL}{9.5} = \frac{1}{7.5}$$

from which $$HL = \frac{9.5}{7.5} = 1.27$$

and $$HK = 2.54 \text{ cm.}$$

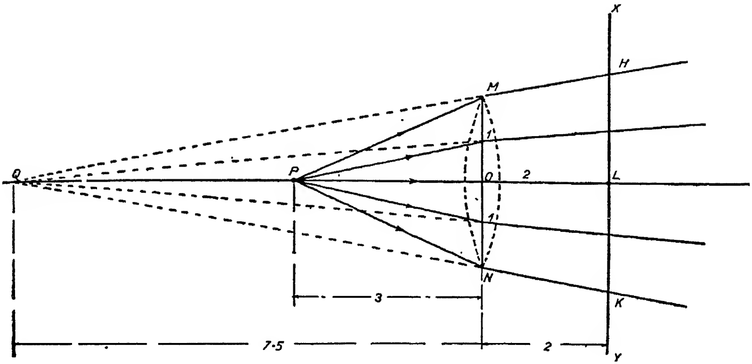


FIG. 60.

NOTE. — Since the lenses we are dealing with are thin, it is desirable in diagrams drawn to scale, such as in Fig. 60, to represent the lens by a straight line.

45. Distant Point Source and Parallel Rays. — Suppose we are looking through a lens at a distant object, for example, a tree 50 ft. high, 100 yds. away. If we represent the tree by the line

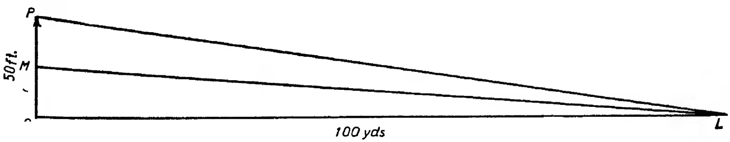


FIG. 61.

PQ, Fig. 61, and make any attempt at drawing a diagram to scale, it is evident that a lens of an inch or two aperture will be represented by an extremely small dot at *L*. Moreover, all the rays coming from a point on the object such as *P*, and incident

on the lens, will be represented by the line PL, all rays from another point M by the line ML whose direction differs from PL.

Suppose, however, we represent the *lens* by the line AB, Fig. 62, and again attempt to make a diagram to scale. It will now be evident that it is impossible to show the distant object, and that we must represent all rays coming from the point P by a bundle of rays which for all practical purposes are parallel; those from Q by another parallel bundle whose direction is different from the one from P. The first bundle will give rise to the point image P', the second to the point Q', and so for every point on

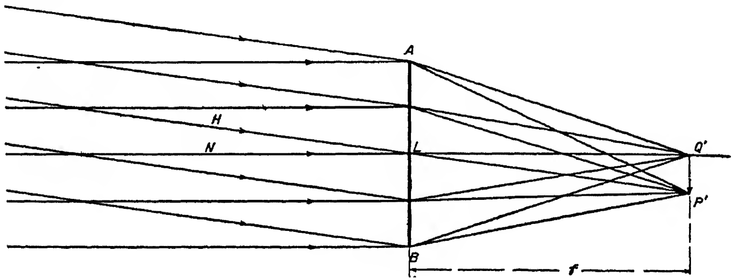


FIG. 62.

the object there will be a corresponding point on the image located in the *focal plane of the lens*.

It will be seen, too, that HLN is the angle subtended at the lens by the object, as well as (since $Q'LP' = HLN$) that subtended by the image. If then the angular size of an object is known, the *linear* size of the image may be readily calculated, since

$$\tan Q'LP' = \frac{P'Q'}{LQ'}$$

or

$$\tan HLN = \frac{P'Q'}{f},$$

or $P'Q'$, the linear size of the image, $= f \times$ (tangent of the angle subtended at lens by object).

In actual practice we have as a rule only to do with angles so small that we may use either the tangent, the sine or the angle expressed in radian measure.

46. Combination of Thin Lenses. — Problems dealing with the passage of light through two or more co-axial lenses may readily be solved either graphically or by calculation, if we take one lens at a time. A concrete example will best illustrate the method.

Two converging lenses, with focal lengths 20 cm. and 30 cm. are placed co-axially 10 cm. apart. Find the position of the image of an object placed 60 cm. from the first lens (*a*) *by calculation*, (*b*) *graphically*.

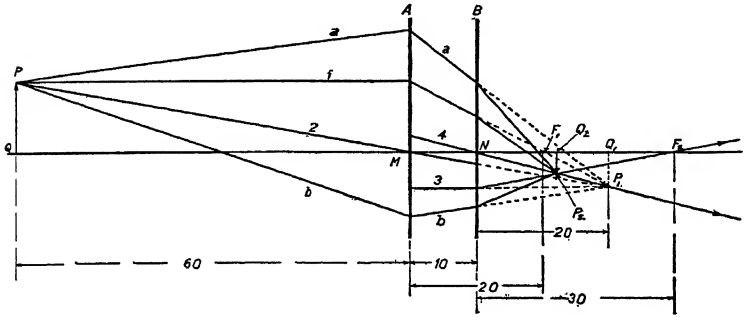


FIG. 63.

(*a*) If we take the first lens, *A*, Fig. 63, we have, using the relation $\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$, and substituting for $p = -60$, $f = +20$,

$$\frac{1}{q} \quad \frac{1}{-60} - \frac{1}{+20},$$

from which $q = +30$ cm.

This means that each cone of light emerging from lens *A* is incident on lens *B* in such a way that the rays would meet at a point 20 cm. beyond *B*, had this lens been absent.

Therefore, dealing with lens *B*, since for it $p = +20$, $f = +30$, we have

$$\frac{1}{q} - \frac{1}{+20} \quad \frac{1}{+30},$$

from which $q = +12$ cm.

(*b*) In solving the problem graphically we first of all represent to scale the positions of the object P, the two lenses and the two principal foci, F_1 and F_2. We then locate the image P_1Q_1 due to lens A, by drawing the principal rays 1 and 2 in the usual way, dotting that portion of the rays beyond B. By next drawing the complete cone of rays from P, we know the path of every ray in this cone in the region between the lenses, since each must fall on lens B in a direction making for the point P_1.

To find the final image all that is now necessary is to take those two rays between the lenses whose paths can be drawn on *emergence* from lens B. Thus ray 3 incident on B parallel to the axis (and making for P_1) emerges through F_2, while ray 4 incident on the centre of B (and also making for P_1) goes through this lens with direction unchanged. P_2, the point of intersection of these two rays, then gives the position of the final image of P.

Similarly for every other point on the object we might trace corresponding bundles, obtaining finally the complete image P_2Q_2. If such a diagram is carefully made to scale, it will be found by actual measurement that the distance NQ_2 is close to 12 cm.

It will readily be seen that both methods may be applied to the case of any number of lenses placed co-axially. The image due to one becomes the object for the next. Care must be exercised, however, to give all distances their proper signs.

47. Thick Lens. — Much the same method is applicable to the case of a thick lens, that is, one whose thickness cannot be neglected in comparison with object and image distances, and with radii of curvature of the surfaces. While in this book we are primarily concerned with physical optics, a brief reference will be made to the thick lens because the treatment we shall give provides another good example of the application of the principles we have been discussing in this and the preceding chapter. Our attention will be limited to the case of a thick lens bounded on both sides by the same (less dense) medium — a glass lens in air, for example.

Our problem then is, given a source such as RS, Fig. 64,

placed p cm. from the first surface of a lens of thickness $AB = t$, with faces of known curvatures, to find where the image is situated. In solving this problem we shall apply the same general method as has been used for the combination of two thin lenses, treating

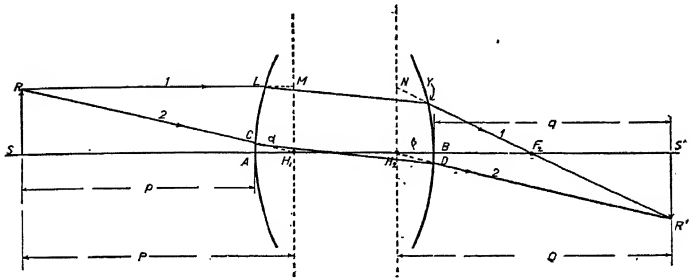


FIG. 64.

the image formed by the first *surface* as the object for the second. Use will be made of the following relations which the student should prove for himself by the method of change of curvature of wave-front, but none of which he should attempt to memorize.

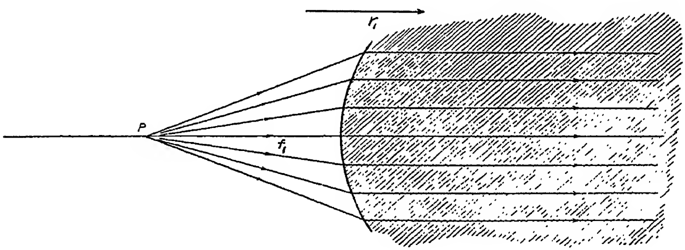


FIG. 65.

(1) If a point source P, Fig. 65, is placed in air at such a distance f_1 from a convex surface separating air from a denser medium of index n that parallel rays traverse this medium, then

$$\frac{1}{f_1} = -\frac{n-1}{r_1}, \tag{4.01}$$

where r_1 = radius of curvature of the surface.

(2) Similarly, if parallel rays (as in Fig. 66), traversing a more dense medium, are incident on a spherical surface, from which

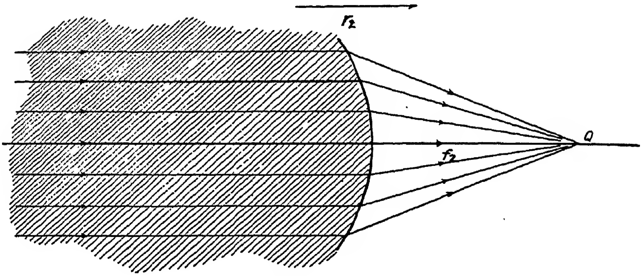


FIG. 66.

they emerge converging to a point Q in air, distant f_2 from the surface, then

$$\frac{1}{f_2} \qquad \frac{n-1}{r_2}: \tag{4.02}$$

where r_2 = radius of curvature of the surface.

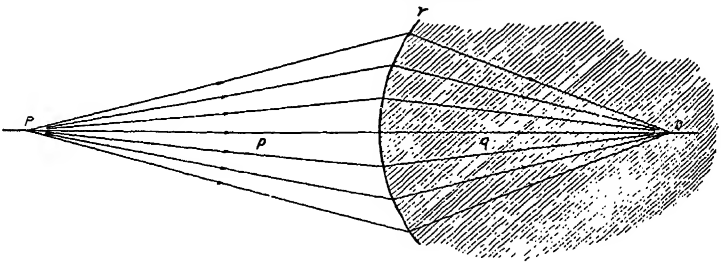


FIG. 67.

(3) If a point source P is placed as in (1) but at a distance not equal to f_1 so that an image Q is formed at a finite distance from the surface, as in Fig. 67, then

$$\frac{n}{q} - \frac{1}{p} = \frac{n-1}{r_1} = -\frac{1}{f_1}. \tag{4.03}$$

(4) Again, if rays which are not parallel strike a refracting surface as in Fig. 68, where they are making for a *virtual object P*

(or when coming from a real object) and a point image is formed in air, as at Q, then

$$\frac{1}{q} - \frac{n}{p} = \frac{n - 1}{r_2}$$

or

$$\frac{1}{q} - \frac{n}{p} = \frac{1}{f_2}. \tag{4.04}$$

(See section 36 for proof.)

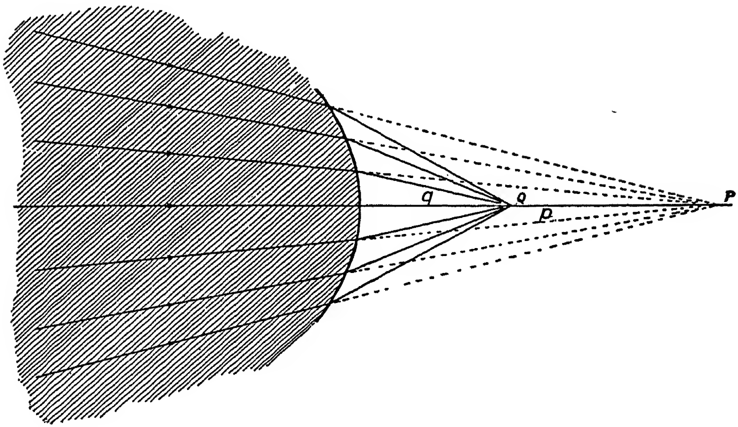


FIG. 68.

Consider now the case of the thick lens represented in Fig. 64. Dealing with refraction *at the first surface*, and using relation (4.03), we have

$$\frac{n}{q'} \cdot \frac{1}{p} \qquad \frac{-1}{r_1} = -\frac{1}{f_1}, \tag{4.05}$$

where q' = the distance of the image from this surface.

Since the image formed by the first surface becomes the object for the second surface, we have, for that surface,

$$p = q' - t \text{ (due regard being paid to sign of } q'),$$

and, therefore, applying relation (4.04),

$$\frac{n}{q' - t} \qquad \frac{n - 1}{r_2} = \frac{1}{f_2}, \tag{4.06}$$

where q is the distance of the final image *from the second surface.*

By combining relations (4.05) and (4.06), eliminating q', we obtain by simple but somewhat laborious algebra

$$\frac{1}{q - \beta} - \frac{1}{p - \alpha} = \frac{1}{F}, \tag{4.07}$$

where for simplicity we have written

$$\alpha = -\frac{f_1 t}{n(f_2 - f_1) - t}, \tag{4.08}$$

$$\frac{f_2 t}{n(f_2 - f_1) - t}, \tag{4.09}$$

$$F = -\frac{n f_1 f_2}{n(f_2 - f_1) - t}. \tag{4.10}$$

Since α, β, and F are quantities depending only on the constants r_1, r_2, n, and t, their values in any given problem can readily be found. Hence relation (4.07) provides us with the desired means of locating the image formed by a thick lens. Its use is illustrated by the following example.

48. *An object is placed 60 cm. from a thick double convex lens, with faces of radii of curvature 20 cm. and 30 cm., with thickness = 2 cm., and index = 1.5. Find the position of the image.*

Here we have

$$f_1 = -\frac{r_1}{n - 1} \qquad \frac{+20}{1.5 - 1} \qquad -40 \text{ cm.},$$

$$f_2 = -\frac{r_2}{n - 1} \qquad \frac{-30}{1.5 - 1} = +60 \text{ cm.},$$

$$F = -\frac{1.5 \times -40 \times +60}{1.5(+60 + 40) - 2} = +24.3 \text{ cm.},$$

$$\alpha = -\frac{-40 \times 2}{1.5(+60 + 40) - 2} = +0.54 \text{ cm.},$$

$$\beta = -\frac{+60 \times 2}{1.5(+60 + 40) - 2} = -0.81 \text{ cm.}$$

Therefore, since $p = -60$,

$$\frac{1}{q - (-0.81)} - \frac{1}{-60 - (+0.54)} = \frac{1}{+24.32},$$

from which $q = +39.85$ cm.

49. Suppose, now, we let

$$Q = q - \beta,$$
$$P = p - \alpha,$$

relation (4.07) then becomes

$$\frac{1}{Q} - \frac{1}{P} = \frac{1}{F}. \tag{4.11}$$

Since the *form* of this relation is identical with that used for a thin lens, we conclude that with a knowledge of α, β, and F the problem of a thick lens is reduced to that of a thin one. Putting it in another way, this means that if we measure the object distance not from the first surface (when it is p) but from a point on the axis distant α from this surface (when it is P), and the image distance, not from the second surface (when it is q), but from a point distant β from this surface (when it is Q), *then the thin lens relation holds*. Such points are called *principal points*, and planes through them perpendicular to the axis *principal planes*.

It follows, moreover, that the graphical method of locating images for thin lenses must hold also for thick lenses, provided we fulfil this condition which makes the same mathematical relation applicable; that is, provided we measure object distances from the first principal plane, image distances from the second, and use F as the focal length. This means that, if we imagine light incident on a thin lens coincident with the first principal plane, and emergent from the lens when coincident with the second principal plane, the usual principal rays may be used. The method will be clear from a re-examination of Fig. 64.

In this diagram RS is an object, the position of whose image as formed by a thick lens we wish to find. Having calculated α, β, and F, we locate on the axis the principal points by measuring from the first surface $AH_1 = \alpha$ (assumed positive in this example),

and from the second surface $H_2B = \beta$ (negative in this case). The point F_2, the principal focus, is then located by measuring $H_2F_2 = F$, while the principal planes are drawn through H_1 and H_2.

It then follows from the above paragraph that a ray 1 leaving the point R parallel to the axis and incident on the first principal plane at M will leave the second principal plane at the point N, where $H_1M = H_2N$, and finally emerge in a direction passing through F_2.

Similarly a ray 2 leaving P in a direction that passes through the point H_1 will leave the lens in a parallel direction which passes

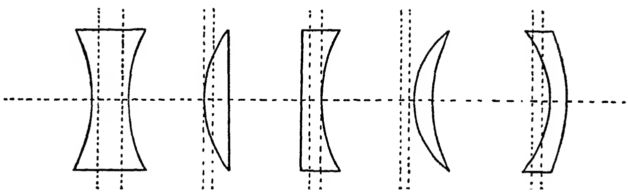


FIG. 69.

through H_2. The point of intersection of these two rays will then give the required position of R', the image of P.

The student should realize that the actual path of these rays *inside* the lens is not along the lines LM, CH, nor NY, H_2D, but rather along the lines LY and CD.

The position of the principal points depends on the shape of the lens. By taking concrete examples it is left to the student to show that for lenses of the standard shapes illustrated in Fig. 69, the principal planes have the positions indicated by the dotted lines. No attempt has been made to draw these diagrams to scale. They indicate, for each lens shape, the approximate position of the planes.

50. Nodal Points. — In the general case of a thick lens bounded by two *different* media *nodal points* are defined as points on the axis such that a ray traversing the first medium in a direction

passing through one point, emerges in the second medium in a parallel direction, which passes through the second point. Obviously, by this definition, H_1 and H_2 in the above diagram (where the medium is the same on both sides of the lens) are nodal as well as principal points. In the general case, however, there are four distinct points, principal points being defined in the following way.

Let NF_2, Fig. 70, be the emergent path of a ray LP incident on a thick lens parallel to the axis, and let M be the point of intersection of these two directions. Then, by taking MH_2 perpendicular to the axis, the position of H_2, the principal point

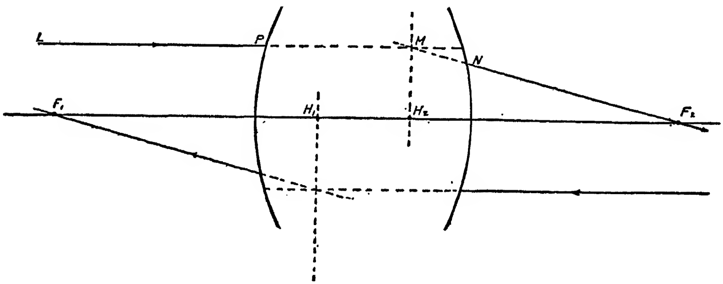


FIG. 70.

related to the second surface, is defined. Similarly by taking a ray parallel to the axis in the opposite direction, we define the other principal point H_1. When the bounding media are different, as already noted, these points are not nodal points. This case, however, it is not our purpose to consider further. We have simply sought to show how, when we have to do with a thick lens made of a material such as glass and surrounded by air, the ideas are essentially similar to those necessary for the study of a thin lens.

50A. Equivalent Focal Length of Two Thin Lenses Not in Contact. — An alternate method to that given in section 46 may

be used to solve problems relating to a combination of two thin lenses separated by a finite distance. The method consists in replacing the two lenses by principal planes, and then, with the aid of an equivalent focal length, applying equation (4.11) and Fig. 64 in almost exactly the same way as has been used in the treatment of a thick lens.

Consider an object p cm. away from a thin lens of focal length f_1 beyond which, at a distance of t cm., there is a second lens of focal length f_2. Then, if q' is the distance of the image formed by the first lens from it, and q is the distance of the final image from the second lens, we can write (remembering that the image due to the first lens is the object for the second)

$$\frac{1}{q'} - \frac{1}{p} = \frac{1}{f_1},$$

and
$$\frac{1}{q} - \frac{1}{q'-t} = \frac{1}{f_2}.$$

Eliminating q' from these relations, we can obtain by straightforward but laborious algebra

$$\frac{1}{q-\beta} - \frac{1}{p-\alpha} = \frac{1}{F},$$

or
$$\frac{1}{Q} - \frac{1}{P} = \frac{1}{F},$$

where
$$F = +\frac{f_1 f_2}{f_1 + f_2 - t}, \tag{4.12}$$

$$\alpha = +\frac{f_1 t}{f_1 + f_2 - t}, \tag{4.13}$$

and
$$\beta = -\frac{f_2 t}{f_1 + f_2 - t}. \tag{4.14}$$

All these relations, it must never be forgotten, are to be used in accordance with the sign convention adopted in this book.

It follows from the above work, that, provided (1) object distances P are measured from the first principal plane, and image

distances Q from the second, and (2) the equivalent focal length F is known, problems relating to a combination of two thin lenses may be solved by the use of the simple lens relation $\frac{1}{Q} - \frac{1}{P} = \frac{1}{F}$.

The method is illustrated by the following solution of the problem already worked in section 46.

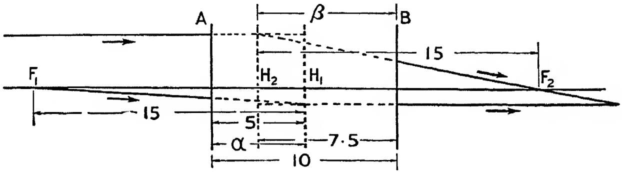


FIG. 70A.

In this problem we are given $f_1 = -60$, $f_2 = +30$, $t = 10$, and $p = -60$. Substituting these values in relations (4.12), (4.13), and (4.14), we obtain

$$F = +\frac{20 \times 30}{20 + 30 - 10} = +15,$$

$$\alpha = +\frac{10 \times 20}{20 + 30 - 10} = +5,$$

$$\beta = -\frac{10 \times 30}{20 + 30 - 10} = -7.5.$$

Using these values, the positions of the principal planes, through H_1 and H_2, and the focal points at F_1 and F_2 have been marked in Fig. 70A.

We can now write $P = p - \alpha = -60 - 5 = -65$. Hence, applying the relation

$$\frac{1}{Q} - \frac{1}{P} = \frac{1}{F},$$

we obtain $$Q = 19.5,$$

and therefore, $$q - \beta = 19.5,$$

or $$q - (-7.5) = 19.5,$$

or $$q = +24.$$

That is, the final image is 24 cm. beyond the second lens as previously found by the other method.

It will be noticed that in Fig. 70A the second principal plane is nearer the first lens than the first principal plane. This not infrequently happens, but need cause no confusion in the mind of the student, because always the object distance P is measured from the first principal plane no matter where it is situated.

The student should examine, in Fig. 70A, the graphical method of locating an image by the use of principal rays and principal planes.

50B. Linear Magnification. — From his work in elementary physics, the student will recall that the linear magnification

$$m = \frac{\text{image length}}{\text{object length}} = \frac{q}{p}.$$

For a combination of two lenses, we may obtain the linear magnification by either of the methods which have been used in the solution of the problem in section 46. Thus, if we use the method of section 46, we write

$$m \text{ for the first lens } = \frac{30}{60},$$

and $$m \text{ for the second lens } = \frac{12}{20},$$

$$\text{from which } m \text{ for the combination } = \frac{30}{60} \cdot \frac{12}{20} \quad 0.3.$$

Making use of the equivalent lens method, we write at once

$$m \text{ for the combination } = \frac{Q}{P} = \frac{19.5}{65} \quad 0.3.$$

From the point of view of Physical Optics, the author of this text considers it more instructive and simpler for the student to adopt the method which takes one lens at a time rather than to use a method which involves memorizing formulas (4.12), (4.13), and (4.14).

Problems

1. A converging thin lens of focal length 5 cm. is stopped down to an aperture of 1 cm. If an eye is placed at a point on the axis 1 cm. from the lens and a scale with millimeter divisions is placed so that its virtual image is formed 200 cm. from the lens, find the number of divisions of the scale which can be seen without moving the eye.

2. An object is placed 3 cm. from a thin lens of focal length 5 cm. and aperture 2 cm. Find the largest size the object can have so that it may be seen by an eye placed at a point on the axis 2 cm. from the lens.

3. A point object is placed (*a*) 100 cm., (*b*) 10 cm. from a converging lens of aperture 5 cm. and focal length 15 cm. Find the greatest distance through which an eye, placed 50 cm. from the lens, can be moved (in a direction normal to the axis) and in all positions see the image.

4. A beam of sunlight falls on a divergent lens of focal length 10 inches; 20 inches beyond this lens is placed a convergent lens of 15 inches focal length. Find where a screen should be placed to receive the final image of the sun.

5. A convergent lens, focal length 10 inches, is placed 12 inches from a gas flame; then 36 inches beyond the first lens is placed a divergent lens of focal length 16 inches. Find the position and size of the final image; is it real or virtual?

6. Two convex lenses of the focal lengths 20 cm. and 30 cm. are 10 cm. apart. Calculate the position and length of the image of an object 20 cm. long, 100 cm. in front of the first lens.

7. Replace the first lens in the above problem by a concave lens of the same focal length and determine the position and magnitude of the image.

8. A luminous object is placed 40 cm. from a converging lens of 20 cm. focal length. 30 cm. beyond the lens is placed a concave mirror of radius of curvature 30 cm. Find (*a*) graphically, (*b*) algebraically the position of the image formed after reflection from the mirror.

9. A small object is placed at a point on the axis 30 cm. from a converging lens (of focal length 15 cm.), 10 cm. beyond which is a second co-axial diverging lens of focal length 10 cm.

(*a*) Find where the image is.

(*b*) If the aperture of the converging lens is 6 cm., that of the diverging lens at least as great, find through what distance an eye may be moved along a line (perpendicular to and intersecting the axis at a point 40 cm. from the centre of the diverging lens) and in all positions see the image.

10. An object is placed 40 cm. from a converging lens (of focal length 20 cm.) 32 cm. beyond which is a diverging lens of focal length 6 cm.

(*a*) Find (1) graphically, (2) by calculation, the position of the final image.

(*b*) If the aperture of the converging lens is 4 cm. and a point source is placed 40 cm. from it, find the smallest aperture of the diverging lens (placed as above) such that all the light leaving the converging lens is transmitted by the diverging.

11. A small object is placed at a point on the axis 40 cm. from a converging lens of focal length 20 cm. and aperture 2 cm., 15 cm. beyond which is a second converging lens of focal length 15 cm. Find

(*a*) the position of the final image;

(*b*) the smallest aperture of the second lens such that *all* the light which passes through the first lens from the small object will also pass through the second.

12. A converging lens of focal length 20 cm. and a diverging lens of focal length 12 cm. are placed 18 cm. apart, and a small source of light placed at a point on the axis in front of the converging lens. Where must the source be placed if parallel rays are to emerge from the diverging lens?

13. Given for a thick converging lens of focal length 20 cm. that $\alpha = +1.6$ and $\beta = 0$; find the position of the image of an object 58.5 cm. from the first face of the lens.

14. Prove that the distance apart of the nodal points of a plano-convex lens $= t \quad (n-1)$ where t = thickness of the lens.

15. Given a thick double convex lens, each of whose faces is of 20 cm. radius; thickness 1 cm.; index = 1.5.

(*a*) Locate the principal points.

(*b*) If an object be placed 50 cm. from the nearer face, find by calculation the position of the image.

(*c*) Solve (*b*) graphically by making a diagram to scale.

(*d*) Treating the lens as thin, solve (*b*) and see what error this leads to in the solution.

16. Given a thick double convex lens, with faces of 20 cm. and 30 cm. radius; thickness = 2 cm.; index = 1.5. Locate the principal points.

17. A thick lens which has for constants $\alpha = +3$, $\beta = -2$, forms an image of a *distant* object 18 cm. from the second surface of the lens. Calculate the position of the image of an object placed 97 cm. from the first surface of the lens.

18. Solve questions 6, 9(*a*), 10(*a*), and 11(*a*) by using principal planes and equivalent focal length.

CHAPTER V

VISION THROUGH AN INSTRUMENT. THE TELESCOPE AND THE MICROSCOPE

51. In the experimental study of the nature of light there are few instruments used more frequently in one way or another than the telescope. It is essential, therefore, that the student of physical optics be thoroughly familiar with the basic principles involved in the construction and use of this instrument. In this chapter these will be discussed in detail.

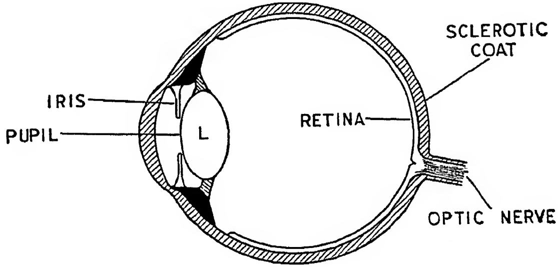


FIG. 71.

Since the general object of looking through a telescope is to enable us to see better than with the unaided eye, it is first of all necessary to understand something about ordinary vision. From the view-point of physical optics the eye is an optical instrument, the essential part of which is the crystalline lens L, Fig. 71, which forms a real image of an object on the retina. The aperture of this lens is fixed by the pupil, the transparent central portion of the opaque iris. This aperture, moreover, is variable, contracting when intense light falls on the eye, dilating for a feeble beam. The image formed on the retina is inverted, just like any other real image formed by a converging lens. As the light which goes to the bottom of the image comes from the top, and as the position of an object "seen" by an individual is in the direction along which

the light actually comes, the interpretation is that of an upright image.

Accommodation. — As everyone knows, it is possible to see distinctly objects over a considerable range of distance. If the reader of these lines suddenly shifts his eyes from the print to an object across the room, he has probably no difficulty in seeing such an object almost as distinctly as the letters on the page of the book. On the other hand, it is not an uncommon experience when one's mind wanders from the page one is reading, to find that the print has become blurred and indistinct. In other words, although the retina is at a fixed distance from the eye lens, it is possible to obtain sharp images of objects at different distances from the eye, and it is equally possible to have the image of an object at a fixed distance either sharply focussed (clearly seen) or indistinct. That this is so in the case of the eye is because of its power of *accommodation*. Muscles are attached to the crystalline lens by means of which the curvature of its faces may be altered, thus altering the focal length.

The question naturally arises, is there any limit to the range over which a normal eye can see objects distinctly? Every-day experience gives the answer, because it is a matter of common observation that, if an object be brought closer and closer to the eye, eventually a position is reached within which clear vision is no longer possible. There is, therefore, a limit to the power of accommodation. A normal eye can see objects distinctly *without overstrain* when they are as near as 25 or 30 cm. Within this limit most people can still see objects clearly to some extent, but in that case the eye is probably overstrained. If persisted in, vision at a distance as close as 20 cm. will lead to eyestrain and headache.

As a small object is brought nearer and nearer the eye from a distance, the image on the retina becomes larger and larger, and *detail* is seen better and better. For purposes of standardization and comparison with vision through an instrument, it is desirable to agree on some standard distance as the normal for unaided vision. Since 25 cm. is about the nearest distance for which clear

vision is normally possible *without overstrain*, this distance is chosen as the standard. It is well to realize again that, while for most people clear vision within this distance is quite possible, it is only so at the risk of eyestrain. The student should note also that the eye muscles are most relaxed when a distant object (or image) is being viewed, that is, when the incident rays from any point on the object (or image) are approximately parallel. For continued vision through an instrument, therefore, it is much easier on the eye if adjustment has been made so that the image is at infinity. (See section 56.)

52. Magnifying Power. — In dealing with the advantages of vision through an instrument, we are concerned with the number

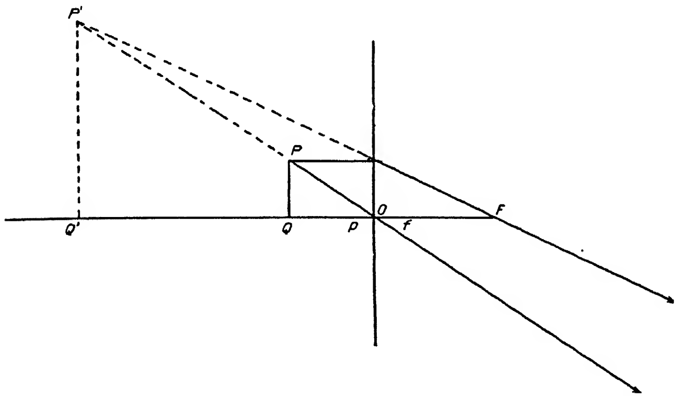


FIG. 72

of times the eye can see better through the instrument than when unaided. In general terms, this number gives us a measure of what is called the *magnifying power* of the instrument. The student must be careful to distinguish clearly between *linear magnification* (the ratio of an *actual* dimension of an image to the corresponding dimension of the object) and magnifying power. If the moon is viewed through a telescope the actual linear diameter of the image may be many times smaller than that of the moon,

while the magnifying power, if vision were aided at all, would have to be greater than unity. The difference should be clear from the following discussion.

If a small object PQ, Fig. 72, placed nearer a converging lens than its focal length, is viewed through such a lens, a magnified virtual image $P'Q'$ is observed. This is the principle of the magnifying glass, or magnifier, or Ramsden eyepiece. To find the magnifying power (M.P.) for this simple aid to vision, our definition must be made more exact. This is done in two ways, both of which will be discussed.

(1) Since the normal distance of distinct unaided vision is 25 cm., one definition frequently used is the following:

$$\text{M.P.} = \frac{\text{linear dimension of image if formed 25 cm. from eye}}{\text{corresponding linear dimension of object}}.$$

The student will at once see that this is nothing but a measure of the linear magnification *when the object is placed in the special position that the image is formed 25 cm. from the eye.* It is this definition which is implied when the microscopist speaks of the magnification of his instrument as so many *diameters.* He is really assuming that the image is formed 25 cm. from the eye, although we shall see presently that the value of the M.P. in this case would not be altered a great deal even if the image were formed at infinity.

We may here state that in the quantitative work which follows, it is assumed, in so far as a single thin lens is concerned, that the eye is placed *immediately* behind the lens and that the error is slight in taking the distance, image to lens, the same as image to eye.

From the above definition, we have, using Fig. 72,

$$\text{M.P.} = \frac{P'Q'}{PQ}$$

$$\frac{Q'O}{QO}$$

$$\frac{25}{p}. \qquad (5.01)$$

Since $\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$, and $q = -25$, this becomes

$$\text{M.P.} = 25\left(\frac{1}{-25} - \frac{1}{f}\right)$$
$$= -\left(1 + \frac{25}{f}\right).$$

Since magnifying power, by definition, is a numerical ratio, we may drop the minus sign and write

$$\text{M.P.} = \left(1 + \frac{25}{f}\right) \tag{5.02}$$

Now in actual practice, when an object is viewed through a magnifying glass, as a rule one has not the faintest idea where the image is, for it is just as possible to view images over a range of distances as objects. If the student will take a short focal length convex lens and try the following simple experiment, he can readily prove this. View a small object through the lens, adjusting until a sharp image is *first* seen (1) by moving the object from a position too close to the lens outwards, (2) by moving from a distant position towards the lens. If the two distances are noted and the focal length of the lens measured, the usual calculation from the relation $\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$ will give the corresponding image distances. The result will show not only the wide range but also that the nearest distance is often considerably less than 25 cm. As in each case the images are equally clear, overstrain and possible headache may be avoided when examining objects through a magnifying glass or an eyepiece, by focussing with the object (or image) far away rather than too near. (In the case of a telescope this means adjusting by moving the eyepiece initially from a position too far out.)

When focussing has been done in this manner, however, the image is approximately at infinity, and in that case the above definition of M.P. cannot be applied. A more general definition is therefore desirable.

(2) It was pointed out above that the nearer the object to the eye, the larger the image on the retina and the better detail can be seen (always assuming we are outside 25 cm.). If, then, when viewing an object through an instrument, the size of the image *on the retina* is increased, vision will be improved and detail better seen. Now a glance at *a* and *b*, Fig. 73, where $P_1Q_1 = P_2Q_2$, will show that the greater the angle subtended at the eye by an

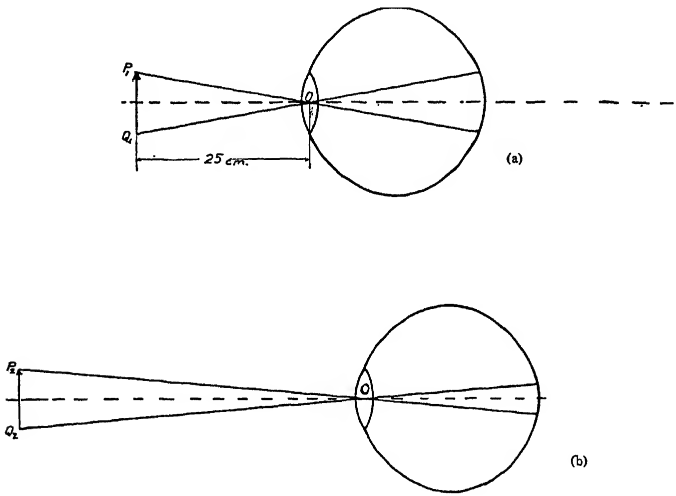


FIG. 73.

object, the greater the size of the image on the retina. We may, therefore, take as a measure of the distinctness of unaided vision, the angle subtended by a distant object at the eye, or, in the case of a small object which may readily be brought near the eye, by the angle subtended when it is at a distance of 25 cm. Hence we have the following more general definition for the magnifying power of a magnifying glass:

$$\text{M.P.} = \frac{\text{angle subtended at eye by image}}{\text{angle subtended at eye by object when placed at 25 cm.}}$$

By this definition we have, again using Fig. 72, and *a*, Fig. 73,

$$\text{M.P.} = \frac{\angle\, P'OQ'\ (\text{Fig. } 72)}{\angle\, P_1OQ_1\ (\text{Fig. } 73a)}.$$

In actual practice the angles with which we have to deal are essentially small (think, for example, of the angle subtended by a letter of print 1 mm. wide at an eye 250 mm. away). We may express those angles, therefore, most simply in radian measure. Thus

$$\angle\, P_1OQ_1\ (\text{Fig. } 73a) = \frac{P_1Q_1}{25}$$

and

$$\begin{aligned} \angle\, P'OQ'\ (\text{Fig. } 72) &= \angle\, POQ, \\ &= \frac{PQ}{p}. \end{aligned}$$

$$\therefore \quad \text{M.P.} = \frac{\dfrac{PQ}{p}}{\dfrac{P_1Q_1}{25}},$$

$$25$$

since P_1Q_1 and PQ each represent the actual object viewed.

It will be seen that this relation is the same as that derived from the other definition. There is a difference, however, for p *is not now confined to the particular value for which the corresponding image distance is 25 cm.* It may be any value, provided the image lies between 25 cm. and infinity, for in all such cases clear vision without overstrain is possible. There is, therefore, a range of values for the M.P. lying between limits corresponding to these two positions of the image.

(*a*) *When the image is at 25 cm.* In this case we have the expression already deduced:

$$\text{M.P.} = \left(1 + \frac{25}{f}\right). \qquad (5.02)$$

(*b*) *When the image is at infinity*. For this to be the case, p must equal f and at once we may write

$$\text{M.P.} = \frac{25}{f}. \tag{5.03}$$

Generally, therefore, the magnifying power of a magnifying glass or a simple microscope or a simple eyepiece lies between these two limits. If f is small, it will be seen that the two differ very little. For example, when $f = 1$ cm., the M.P. ranges from 25 to 26.

53. The Simple Telescope. — The telescope, an instrument used to observe a *distant* object, consists essentially of, (*a*) the *objective*, a converging lens which forms a real image of the distant

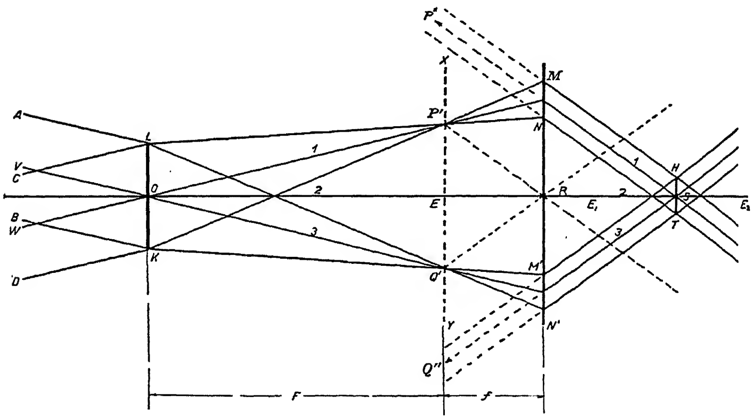


FIG. 74.

object in its focal plane, (*b*) an *eyepiece*, which, in the simplest case, is a single converging lens through which this real image is viewed. The eye-lens, then, is used as a magnifying glass forming a virtual image at any distance between 25 cm. and infinity. Since, as already noted, the eye muscles are subject to the least strain when parallel rays fall on the eye, in the work that follows we shall assume that the instrument has been adjusted so that this is the case.

Fig. 74 shows the optical paths of two bundles through such an instrument, (1) AVB, parallel rays, from a point on one side of a distant object, (2) CWD, another parallel bundle from a point on the other side of the object. Corresponding real point images are then formed by the objective at P' and Q', so that, if a screen (or photographic plate) is placed in the plane $P'Q'$, a sharp image of the distant object is seen (or photographed). Actually the rays spread out from P' and Q', eventually, on emergence from the eye-lens, giving rise to the bundles $MNHT$ and $M'N'HT$. These emergent bundles will consist of parallel rays, if $ER = f$, the focal length of the eye-lens. When these emergent bundles fall on the eye, point images are formed on the retina and the eye sees the final virtual image at $P''Q''$.

Magnifying Power. — Since a large distant object obviously cannot be brought to the position of most distinct vision, the definition of M.P. applicable to a telescope is slightly different from that given above. According to the new definition,

$$\text{M.P.} = \frac{\text{angular size of image}}{\text{angular size of object}},$$
$$= \frac{\angle P'RQ'}{\angle VOW},$$

since $P'R$ is parallel to rays apparently coming from a point on one side of the image, and $Q'R$ parallel to those apparently coming from a point on the other side. We have then,

$$\angle P'RQ' = \frac{P'Q'}{f}$$

again remembering that this angle is much smaller than shown in the figure.

Also

$$\angle VOW = \angle P'OQ',$$
$$P'Q'$$

where $F = OE$, the focal length of the objective.

$$\therefore \quad \text{M.P.} = \frac{F}{f}. \tag{5.04}$$

54. The Eye-Ring. Eye-Point. — It will be observed that the emergent parallel bundles from the two point sources represented in Fig. 74 pass through the area HT. This area is called the *exit-pupil* or the *eye-ring*, and S, the point where the area HT intersects the axis of the telescope, is called the *eye-point*. A little consideration will show that through this area pass not only these two bundles but also those coming originally from every other point on the object which can be seen through the instrument. The exit-pupil, therefore, defines the place at which an eye must be placed in order to see the biggest range of object. If placed immediately behind R, the centre of the eye-lens, for example, or at some point E_2, beyond S, the outer bundles shown in the figure would not fall on the eye at all. In such positions the corresponding outer points would then not be seen.

Diameter of Eye-Ring. — For visual observation there is no gain in brightness once the pupil of the eye is filled with light. Thus if HT is greater than the diameter of the pupil, the outermost rays of each individual bundle cannot enter the eye, and, as far as vision is concerned, are useless. It is important, therefore, to know the diameter of the eye-ring. That it becomes greater, the greater the aperture of the objective, should be evident by a glance at the diagram. To find its exact magnitude, we proceed as follows:

$$\begin{aligned} HT &= M'N' = MN \\ &= \theta \times f, \end{aligned}$$

where $\theta = \angle M'Q'N'$ or $\angle MP'N$, and $f = ER =$ focal length of eye-lens. This relation holds only provided θ is small and provided the bundles passing through the telescope are inclined at a small angle to its axis. In actual practice the error is not great in assuming that these conditions are fulfilled, as the student can readily verify for himself by taking suitable approximate dimensions for a concrete case.

But $$\angle MP'N = \angle LP'K,$$

and $$\angle M'Q'N' = \angle LQ'K,$$

$$\therefore \quad \theta = \frac{a}{F},$$

where $a = LK =$ aperture of the objective,
and $F = OE$, the focal length of the objective.

We have, therefore,

$$HT = \frac{a}{F} \cdot f,$$

or, diameter of eye-ring, $$\frac{a}{\text{M.P.}} \cdot \qquad (5.05)$$

We see, then, that HT, the diameter of the eye-ring, varies directly as the aperture of the objective and indirectly as the M.P. of the telescope.

Now it will be seen when we study diffraction that it is desirable to have objectives with big apertures (to enable distant objects very close together to be seen). The consequent increase in the diameter of the eye-ring may be offset by increasing the M.P. of the telescope, that is, by having an objective of long focal length as well as an eyepiece of small focal length. The use of objectives with wide apertures is a great advantage, because the bigger the aperture, the greater the light collected by the instrument and the brighter the image. For visual work, this is only the case provided all the light can enter the eye, but, as we have just seen, this condition can be fulfilled by using high magnifying power.

Position of Eye-Point. — An examination of Fig. 74 will show that rays 1, 2, and 3, all passing through O, the centre of the objective, after emergence from the eye-lens, intersect at S, the eye-point. This point, therefore, must coincide with the position of the image of the centre of the objective formed by the eye-lens. Since OR, the object distance $= F + f$, and since the focal length f is known, the corresponding image distance may readily be found from the usual lens relation. Calculation will show that S lies at a distance from the eye-lens a little greater than its focal length.

Field of View. — In the first paragraph of this section it has been pointed out that the maximum *field of view* is obtained when the eye is placed at the eye-point. When the eye is in this position, the actual extent of the field of view depends on the diameter

of the eye-lens. In Fig. 74, for example, unless the aperture of this lens is NRM', the outermost bundles from an object of the angular width shown in that figure could not be transmitted through the eye-lens. Obviously the smaller the aperture of this lens, the less the angular width of the largest object which can be seen.

Because of defects in simple lenses, a question to which further reference is made in the next section, it is frequently not wise to use the full aperture of the eye-lens, and the field of view is still further restricted by a *stop*, or diaphragm. In the simple telescope, as illustrated in Fig. 74 or Fig. 76, this should be placed in the plane $P'Q'$ where the real image is formed by the objective. This diaphragm conveniently carries the cross-hair lines. (See section 56.)

55. The Eyepiece. — When a single lens is used as the eye-lens, there are several objections to placing the eye at the eye-point, where, as we have just seen, maximum field of view is obtained. We note three.

(1) It is desirable to have the eye immediately behind the eye-lens, at E_1, Fig. 74, for example, rather than at S, so that it may be shielded from light not coming through the telescope.

(2) When the eye is at the eye-point, it sees the outermost portions of the image by means of little bundles of rays which have passed through the edges of the eye-lens, the central portion by means of rays through the centre of the lens. Now the student will recall the phenomenon of spherical aberration, the fact that, for apertures of any considerable width, the focal length of the central portion of a single lens differs from that of the outer portions. It follows that vision through such a lens, with the eye at the eye-point, will give rise to an image, not all portions of which are at the same distance from the eye. In such a case, therefore, a good clear image over the whole field of view is not possible. As we have noted above, this defect may be remedied by using a stop to reduce the aperture, but the remedy involves cutting down the field of view.

(3) In the next chapter it will be shown that the focal length of a single lens varies with the colour (wave-length), a defect called *chromatic aberration*. With a single eye-lens, therefore, particularly one of wide aperture, the image may show a certain amount of colour.

Because of these objections, in a good practical telescope the above simple telescope is somewhat modified. In the first place, the objective, instead of being a single lens, is an *achromatic combination* consisting of a converging and a diverging lens, in close contact, each made of a different kind of glass. As such a combination functions exactly in the manner of the objective shown in Fig. 74, no further reference need here be made. (See, however, section 102.)

In the second place, the eye-lens is replaced by a combination of two lenses, called an *eyepiece*. There are several kinds of eyepieces, but in this section we shall confine our discussion to the Ramsden, the one in common use in telescopes attached to such instruments as the spectrometer (see section 61). When cross-hairs are necessary, this is the usual type used.

The Ramsden eyepiece consists of a combination of two converging lenses, of equal focal length, separated by a distance equal to two-thirds their common focal length. The lens nearest the eye we shall call the eye-lens, that nearest the object, the field lens. While this eyepiece functions in exactly the same way as the single eye-lens, it is an improvement on it because the defects enumerated above are much less marked. To begin with, the eye-ring is now much nearer the eye-lens. In locating its position we shall again assume that the instrument is adjusted, as in the previous case, so that parallel bundles of rays leave the telescope. We must first of all, therefore, find where the eyepiece should be placed with reference to the image $P'Q'$ formed by the objective, so that parallel bundles of rays emerge.

This is essentially a problem of the type worked in section 46 and may be solved by either of the two methods there given.

(*a*) By calculation. It is left as an exercise for the student to prove that if p = distance of an object from lens A, Fig. 75,

where A and B represent the two lenses of a Ramsden eyepiece from which parallel rays emerge

$$p = -\frac{f}{4}.$$

The real image $P'Q'$, therefore, should be placed at this distance from the field lens of the eyepiece.

(b) It is more instructive, however, to solve the problem graphically. Suppose, again, that a bundle of parallel rays leaves the lens B, Fig. 75. We may locate the exact position of P_1 by taking on the central ray 1 the point which is distant f cm. from the lens B. Once P_1 is located the direction of the complete bundle incident on B may then be drawn as shown in the figure.

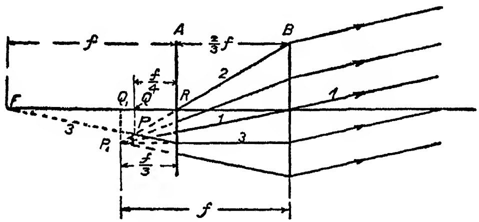


FIG. 75.

Now lens A is present and the position of P_1 enables us, as yet, to draw the actual rays only in between the lenses. Since, however, we know that the principal ray 2 emerges from A through its centre unchanged in direction, and that the principal ray 3 emerging from A parallel to the axis (and apparently from P_1) must have been incident on A in a direction passing through F (a point f cm. from R), P the point of intersection of these two rays gives us the original point from which the parallel rays must have come. If such a diagram is made to scale, the student will find by actual measurement that the distance $QR = \frac{f}{4}$.

When, therefore, we replace the simple eye-lens of Fig. 74 by a Ramsden eyepiece, the image $P'Q'$, formed by the objective, must be at a distance $\frac{f}{4}$ from the field lens. The diagram for

such a case is given in Fig. 76, in which the focal length of the objective has been taken three times f. In tracing the outermost bundles through the eyepiece, one lens has been taken at a time and the usual principal rays used. Thus, taking lens A, P'' has been located as for any converging lens, by using rays 1 and 2, 1 through R the centre of A, 2 incident parallel to the axis, emerging in a direction passing through F_2, where $RF_2 = f$. When P'' has been found, it is then possible to draw the actual path of the shaded bundle in between the lenses, since all such rays must emerge from A as if coming from P''. Finally, the direction of the bundle

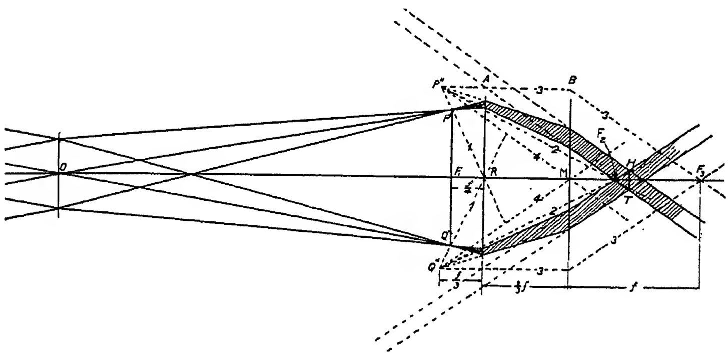


FIG. 76.

emerging from B is found by making use of the principal rays 3 and 4. Ray 4 incident on the centre of B as if coming from P'' goes through unchanged in direction, while ray 3 incident on B, parallel to the axis, but again as if from P'', emerges in a direction passing through F_3, where $MF_3 = f$.

If the diagram is made carefully, 3 and 4 on emergence will be parallel, thus showing that the final image is at infinity, as well as giving the direction of the actual emergent bundle of parallel rays (shaded in the figure). In Fig. 76, all rays which do not actually exist have been indicated by dotted lines.

In a similar manner, the path of a bundle of rays from a point on the other side of the object may be drawn, and hence the

position of the eye-ring HT located. It will be seen by comparing Fig. 74 with Fig. 76, that in the case of a telescope equipped with a Ramsden eyepiece, the eye-ring is considerably closer to the instrument than when a single eye-lens is used. This may be proved by direct calculation, since the eye-point coincides with the image of O, the centre of the objective, formed by the eye-piece. It is left as an exercise for the student to make the necessary calculation. The first objection to the simple arrangement noted at the beginning of this section has, therefore, to a considerable extent been overcome by the use of an eyepiece.

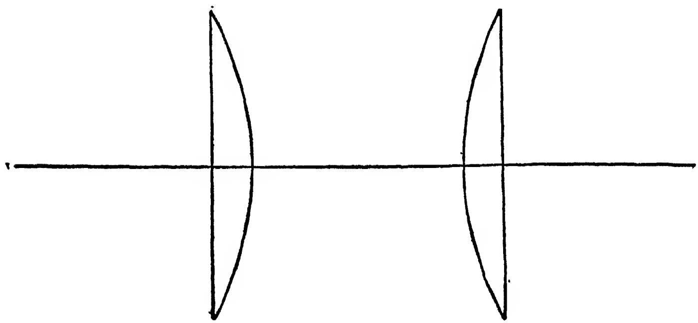


FIG. 77.

Defects due to spherical and to chromatic aberration are also lessened by the use of such a combination of two lenses. Into the details of the improvement for spherical aberration we shall not enter, except to point out that by using two plano-convex lenses placed as in Fig. 77, the total deviation of a ray through the system is the sum of small deviations at each of four surfaces. By having the deviation at each surface of a lens nearly the same spherical aberration is lessened. (Note again section 40.) In the next chapter achromatic lens combinations will be considered.

Magnifying Power. — The M.P. of a telescope equipped with a Ramsden eyepiece is increased somewhat, as will be evident from the following proof. Since the outermost parallel bundles emerging from the telescope are parallel to $P''M$ and $Q''M$ re-

spectively, the angle $P''MQ''$ is the angle subtended at the eye by the image. Therefore,

$$\text{M.P.} = \frac{\angle P''MQ''}{\angle P'OQ'}$$

$$= \frac{\dfrac{P''Q''}{f}}{\dfrac{P'Q'}{F}},$$

since $P''Q''$ is formed at a distance f from lens B.

$$\therefore \quad \text{M.P.} = \frac{P''Q''}{P'Q'} \cdot \frac{F}{f}.$$

But
$$\frac{P''Q''}{P'Q'} = \frac{\dfrac{f}{3}}{\dfrac{f}{4}} = \frac{4}{3}$$

and hence,
$$\text{M.P.} = \frac{4}{3} \cdot \frac{F}{f}.$$

Corollary: If a Ramsden eyepiece is used as a magnifier, it follows that it is equivalent in M.P. to a single lens of focal length $\frac{3}{4}f$.

Alternative Treatment. — Problems relating to the Ramsden eyepiece may be solved readily by the use of principal planes and the equivalent focal length. Using the relations (4.12), (4.13), (4.14), given in section 50A, we obtain at once

$$F = +\frac{f_1 f_2}{f_1 + f_2 - t} = \frac{f^2}{2f - \frac{2}{3}f} = +\frac{3f}{4},$$

$$\alpha = +\frac{f_1 t}{f_1 + f_2 - t} = +\frac{f}{2},$$

and
$$\beta = -\frac{f_2 t}{f_1 + f_2 - t} = -\frac{f}{2}.$$

With the aid of these values, in Fig. 77A the positions of the principal points H_1 and H_2 and the focal points F_1 and F_2 have been marked.

A glance at this diagram shows that the distance of the first focal point from the field lens of the eyepiece is given by

$$F_1A = F_1H_1 - \alpha = \frac{3f}{4} - \frac{f}{2} = \frac{f}{4},$$

the result obtained previously by the method illustrated by Fig. 75. Note, also, that BF_2, the distance of the eye-lens to the second focal point, is also $\frac{f}{4}$.

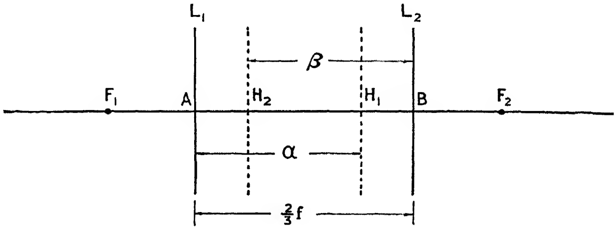


FIG. 77A.

Since the equivalent focal length of the eyepiece is $\frac{3}{4}f$, it is not difficult to prove that

$$\text{M.P. of telescope} = \frac{F}{\frac{3}{4}f} = \frac{4F}{3f}.$$

It is interesting to note that if the eye is placed at F_2, the second focal point of the Ramsden eyepiece — and in actual practice it would not be far from this point — its magnifying power is $\frac{25}{\text{equivalent focal length}}$ *regardless of where the image is.*

To prove this consider a small object, such as PQ, Fig. 77B, viewed through a Ramsden eyepiece used as a magnifier. A ray PN, parallel to the axis and striking the first principal plane at N, emerges from the combination in the direction MF_2, M being the point on the second principal plane corresponding to N on the first principal plane, and F_2 the second focal point. If P is the top of the object, then this ray through F_2 leaves the eyepiece as

if coming from the top of the image. Hence, if the eye is at F_2, the angle MF_2H_2 is the angle subtended at the eye by the image.

$$\therefore \text{ M.P. of the eyepiece} = \frac{\text{angular size of image}}{\text{angular size of object at 25 cm.}}$$

$$= \frac{MF_2H_2}{\dfrac{PQ}{25}}.$$

But
$$MH_2 = NH_1 = PQ,$$
$$\angle MF_2H_2 = \frac{MH_2}{H_2F_2} = \frac{PQ}{\frac{3}{4}f}$$

Hence
$$\text{M.P.} = \frac{\dfrac{PQ}{\frac{3}{4}f}}{\dfrac{PQ}{25}} = \frac{25}{\frac{3}{4}f}.$$

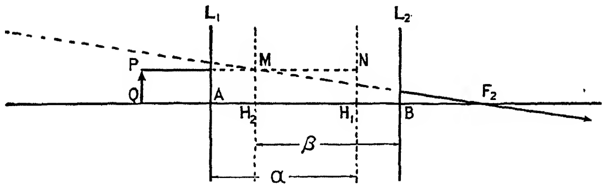


FIG. 77B.

This proof, it will be noted, applies whether the image is at infinity or not, provided, of course, that it is not so near that the eye has to over-accommodate in order to see it.

Huygens' Eyepiece. — This is another type frequently used, especially in instruments where cross-hair lines are not necessary. It consists of two converging lenses separated by a distance equal to one-half the sum of the focal lengths, a condition which ensures that the focal length of the combination is achromatic. The focal length f_1 of the field lens is always greater than f_2, the focal length of the eye-lens, but the ratio $\frac{f_1}{f_2}$ is not always the same. The dia-

grams in Figs. 77C and 77D are made for a ratio of 3 to 1, the highest value used.

With the aid of either of these figures, the student should have no difficulty in proving that, when parallel rays emerge from a Huygens' eyepiece, rays incident on the field lens must make for a virtual point (Q_1 in Fig. 77C, F_1 in Fig. 77D) which is distant $\frac{3f}{2}$ from this lens.

Fig. 77C shows the graphical solution when use is made of the same method as is illustrated in Fig. 75 for a Ramsden eyepiece.

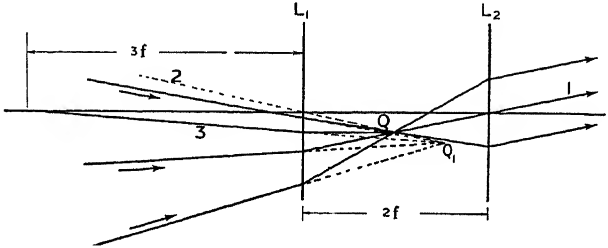


FIG. 77C.

Fig. 77D illustrates the solution when principal planes and the equivalent focal length are used. The student can easily verify that $\alpha = AH_1 = 3f$; $\beta = H_2B = -f$; and the equivalent focal length $= F_1H_1 = H_2F_2 = \frac{3f}{2}$. Taking the beam of parallel rays represented by 1, 2, and 3, we see that, since rays 1 and 2 leave the second principal plane at M and M_1, they must have been (virtually) incident on the first principal plane at N and N_1. Since F_1 is the first focal point, F_1N and F_1N_1 are the virtual paths of these incident rays, and hence C and D must be the actual points of incidence on the field lens. PC and RD (whose directions produced run through F_1) are, therefore, the actual paths of the incident rays which on emergence give rise to rays 1 and 2.

Two things will be evident from either Fig. 77C or Fig. 77D. (1) A Huygens' eyepiece cannot be used as a magnifier of a real object. (ii) If this eyepiece is used in a telescope, it must be so

placed with respect to the objective that the image formed by the objective would, in the absence of the eyepiece, be at Q_1 in Fig. 77C, or F_1 in 77D.

56. Cross-Hair Lines. — It is frequently necessary to set a telescope in an exact position with reference to the real image formed by the objective. When this is the case, cross-hair lines or cross-wires are used, filaments of a spider's web being useful for this purpose. When the telescope with a Ramsden's eyepiece

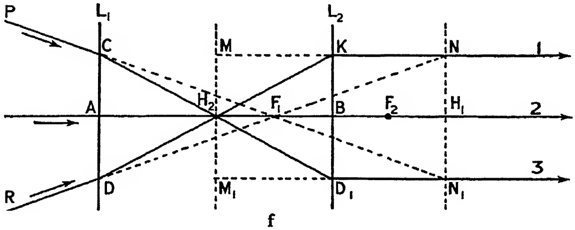


FIG. 77D.

is properly adjusted the cross-wires should coincide with the real image $P'Q'$ (Fig. 74 or Fig. 76). They are attached, therefore, to a diaphragm placed in the barrel of the telescope a short distance from the field lens of the eyepiece. In adjusting the instrument, first of all the eyepiece is moved with reference to the hair lines (preferably from a position too far out) until a sharp image of the cross-wires is obtained. Then, by means of the main focussing screw of the telescope, the system, cross-hairs plus eyepiece, is moved with reference to the objective until a clear image of the distant object is obtained. (This operation, too, should be done starting with the eyepiece too far out. See section 51.) When proper adjustment has been made, there should be no parallax between the image of the cross-hairs and that of the distant object.

If a telescope is equipped with a Huygens' eyepiece and it is necessary to use cross-hairs, these would have to be placed between the field and the eye-lenses in a position coinciding with the real image at Q, Fig. 77C, or H_2, Fig. 77D.

56A. The Microscope. — The magnifying powers of magnifiers are of the order of 10 or 20. When much higher powers are wanted, use is made of the microscope. With the aid of Fig. 77E, the student will recall the elementary principles of this instrument. The objective L_1, a combination of lenses functioning as a single converging lens, forms at P_1Q_1 an enlarged real image of the small object PQ. A magnified virtual image of this real image is then viewed through an eyepiece or magnifier L_2, just as in the

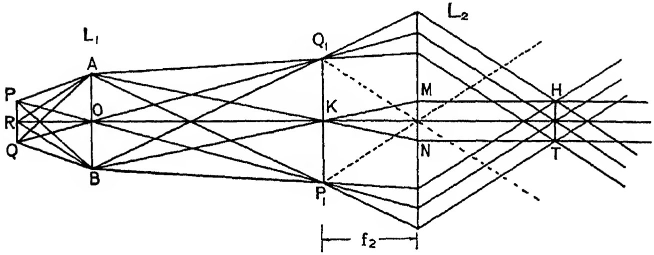


FIG. 77E.

telescope. When the magnification is moderate, the Huygens' eyepiece is frequently the type used. To save the eye, adjustment is preferably made with the final image at infinity.

If m is the *initial magnification*, that is, the linear magnification of the image P_1Q_1,

$$\text{M.P. of microscope} = m \cdot \frac{25}{f_2},$$

where f_2 is the equivalent focal length of the eyepiece. Since the value of the second factor ranges approximately from 5 to 20, f_2 does not exceed a few centimetres.

If a combination of lenses is replaced by the equivalent principal planes, the value of m, which we have previously taken as $\dfrac{\text{image distance}}{\text{object distance}}$, may be written $\dfrac{L}{f_1}$*. In this expression L as

$$* \; m = \frac{P_1Q_1}{PQ} \quad \frac{P_1Q_1}{AH_1} = \frac{P_1Q_1}{BH_2} = \frac{F_2Q_1}{H_2F_2} = \frac{L}{f_1}.$$

shown in Fig. 77F, is the distance of the image P_1Q_1 from F_2, the second focal point of the objective, and $f_1 = H_2F_2$, the focal length of the objective.* Since it is a common practice among manufac-

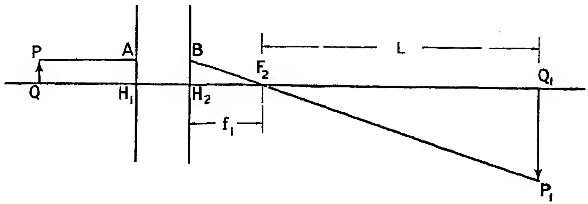


FIG. 77F.

turers to design the length of microscope tubes so that $L = 18$ cm., we may write

$$\begin{aligned} \text{M.P.} &= m \cdot \frac{25}{f_2} \\ &= \frac{L}{f_1} \cdot \frac{25}{f_2} \\ &= \frac{18 \times 25}{f_1 f_2} = \frac{450}{f_1 f_2}. \end{aligned}$$

As it is a common practice to mark the value of m on an objective and the power $\left(\frac{25}{f_2}\right)$ on the eyepiece, the M.P. of the microscope is at once found by multiplying these two values together.

The objective is an achromatic combination of a number of lenses, sometimes as many as ten, so designed that correction is made for chromatism, spherical aberration, and other defects. Fig. 77G, for example, shows the component parts of a high power objective of focal length 1.8 mm. A firm like Bausch and Lomb, of Rochester, N. Y., lists objectives whose focal lengths range from less than 2 mm. to as high a value as 50 mm., with the initial magnification ranging from 2 to almost 100. A good microscope is invariably provided with several objectives of varying powers and two or more eyepieces.

* If the medium is not the same on each side of the objective the focal length on the object side is different from that on the image side, and f_1 is then the value of the latter.

There are several important types of objectives in common use. The one most frequently used is probably the *achromat*, a system of glass lenses made achromatic for two wave-lengths. "The images formed are excellent in every respect, except that some colour is present which becomes disturbing when high power

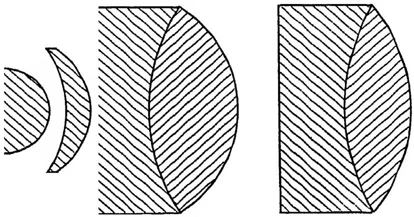


FIG. 77G.

eyepieces and oblique illumination are resorted to." (Bausch and Lomb.) In higher grade achromats, chromatic correction is improved by replacing some of the component glass lenses by fluorite.

The *apochromat* is a system made achromatic for three wave-lengths and in general is more perfectly corrected so that "much higher power eyepieces with subsequent higher total magnification than is recommended for achromatic objectives can be used to good advantage." (Bausch and Lomb.) This type has also a higher numerical aperture (N.A.), for the same initial magnification, than the achromat, an advantage which will appear later.

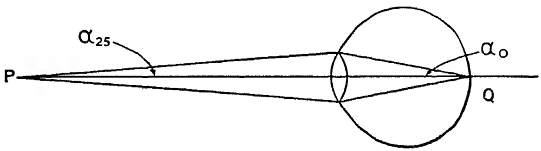


FIG. 77H.

When an *oil-immersion* objective is used, an oil, whose index of refraction is the same as that of the lowest lens of the objective, is the medium between the object and the objective. The advantage of this type is explained in section 103A.

Numerical Aperture. — The magnifying power of a microscope may be expressed in terms of a very important quantity called the *numerical aperture* (N.A.). Although N.A. is important chiefly in connection with resolving power (section 103A), it may conveniently be introduced at this stage by showing its relation to magnifying power. To do so, (1) we make use of the sine

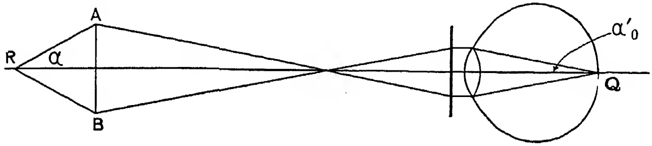


FIG. 77I.

condition $n_1 y_1 \sin \alpha_1 = n_2 y_2 \sin \alpha_2$ (see, again, section 36A); and (2) without in any way altering the correctness of the work given previously, we re-define M.P. as follows:

$$\text{M.P.} = \frac{\text{linear size of image of a small object formed on retina by microscope}}{\text{linear size of image on retina when the small object is viewed directly at a distance of 25 cm. from the eye}}.$$

Let y_0 = linear size of image on the retina when a small object of linear size y is viewed at 25 cm. from the eye;

y'_0 = linear size of image on the retina when the same object is viewed through a microscope;

α_{25} = the object slope angle of a ray from the axial point P, Fig. 77H, of the small object 25 cm. from the eye, to the periphery of the pupil of the eye;

α_0 = the image slope angle of the corresponding ray from the periphery to the axial image Q, Fig. 77H;

α = the slope angle of a ray to the periphery of the objective, from the axial point R, Fig. 77I, when the small object is viewed through the microscope; and

α'_0 = the image slope angle of the corresponding ray to the image Q, Fig. 77I, on the retina.

Applying the sine condition to Fig. 77H, we can write

$$n\, y \sin \alpha_{25} = n_0 y_0 \sin \alpha_0,$$

where n_0 is the index of the medium within the eye. Since the object, when viewed directly, is invariably in air, $n = 1$, and hence

$$y \sin \alpha_{25} = n_0 y_0 \sin \alpha_0. \quad (a)$$

Similarly the sine condition applied to Fig. 77I gives us

$$n\, y \sin \alpha = n_0 y'_0 \sin \alpha'_0 \quad (b)$$

Combining (*a*) and (*b*), and recalling the definition of M.P. given at the beginning of this section, we have

$$\begin{aligned} \text{M.P.} &= \frac{y'_0}{y_0} = \frac{n\, y \sin \alpha}{y \sin \alpha_{25}} \; \frac{\sin \alpha_0}{\sin \alpha'_0} \\ &= \frac{n \sin \alpha \sin \alpha_0}{\sin \alpha_{25} \sin \alpha'_0}. \qquad (c) \end{aligned}$$

The quantity $n \sin \alpha$, that is, the product of the index of refraction of the object medium (generally air but not always) multiplied by the slope angle of the outermost ray from an axial point on the object, is called the numerical aperture (N.A.) of the objective of a microscope. Its value ranges from less than 0.1 to as high as 1.4.

Normal Magnifying Power. — If the optical system is such that the pupil of the eye is completely filled with light by the beam emerging from the eyepiece, the magnifying power is said to be *normal.* For all practical purposes,* this means that $\alpha_0 = \alpha'_0$, from which it follows that

$$\text{normal M.P.} = \frac{\text{N.A.}}{\sin \alpha_{25}}. \quad (d)$$

But α_{25} is the slope angle of a ray from an axial point 25 cm. from the eye to the periphery of the pupil. If, therefore, we take 2.5 mm. as a reasonable value for the diameter of the pupil when the eye is examining a small object in average illumination, then

$$\alpha_{25} = \frac{1.25}{250} = \frac{1}{200}, \text{ from which we have}$$

$$\text{normal M.P.} = 200 \text{ N.A.}$$

* The aperture of the pupil varies with the intensity of the light.

This expression, it should be clearly understood, does not give the actual M.P. of a microscope. It is a useful approximation when the beam emerging from the eyepiece is wide enough to fill the pupil of the eye. The actual diameter of this beam depends on the size of the exit-pupil of the microscope.

Radius of Exit-Pupil. — By referring back to the discussion of the telescope in section 54, the student should have no difficulty in seeing (1) that HT, in Fig. 77E, represents the exit-pupil of the microscope; and (2) that HT is the image of AB, the aperture of the objective, formed by the eye-lens L_2. As in the telescope, so in the microscope, the eye should be placed at the exit-pupil to see the greatest range of object. Microscopes are designed so that the exit-pupil is at a convenient distance, 10 mm. or so from the upper surface of the eyepiece.

Frequently magnifying powers much greater than normal are used. When this is the case, the diameter of the exit-pupil is less than that of the pupil of the eye, as the following proof shows.

If r = radius of exit-pupil, then, in Fig. 77E,

$$2r = HT = MN$$
$$= 2f_2 \tan \phi,$$

where f_2 is the focal length of the eye-lens, and $\phi = \angle AKO$.

Since $\angle ARO = \alpha$ is the object slope angle for the objective, and ϕ is the corresponding image slope angle, then by the sine condition

$$n\, y \sin \alpha = n_1 y_1 \sin \phi,$$

where y_1 is the linear length of the image formed by the objective of a small axial object of corresponding length y.

It follows that m, the initial magnification, is given by

$$m = \frac{y_1}{y} = \frac{n \sin \alpha}{\sin \phi} = \frac{\text{N.A.}}{\sin \phi}.$$

Since OK is many times greater than AO, the angle ϕ is small, hence, with little error $\sin \phi = \tan \phi$, and

$$m \quad \frac{\text{N.A.}}{\tan \phi}.$$

But M.P., the magnifying power $= m \dfrac{25}{f_2}$

$$= \frac{\text{N.A}}{\tan\phi} \cdot \frac{25}{f_2}$$

$$= \frac{25\ \text{N.A.}}{r},$$

or $\qquad r$, the radius of the eye-ring $= \dfrac{25\ \text{N.A.}}{\text{M.P.}}$ cm.

Let us examine three cases.

(*a*) $\qquad$ M.P. = normal

$$= 200\ \text{N.A.}$$

Then

$$r = \frac{25\ \text{N.A.}}{200\ \text{N.A.}}\ \text{cm.}$$

$$= 1.25\ \text{mm.},$$

which is the radius of the pupil of the eye used above, as of course it should be.

(*b*) $\qquad$ M.P. > normal.

Suppose, for example, a M.P. of 500 and a N.A. of 1 is used. Then

$$r = \frac{25\ \text{N.A.}}{500}\ \text{cm.}$$

$$= \frac{250}{500}\ \text{or}\ 0.5\ \text{mm.},$$

a value which is considerably less than the radius of the pupil. The higher magnifying power has been gained, therefore, at the expense of light. This, however, is not a serious matter, because it is generally possible to compensate for the loss of light due to reduced diameter of the beam entering the eye by increasing the illumination.

The advantage of M.P. higher than normal is discussed when the question of the resolving power of a microscope is considered in section 103A.

(*c*) M.P. $<$ normal.

In this case the radius of the exit-pupil exceeds that of the eye-pupil, and the whole beam of light is not utilized. This means that, although the actual brightness of the image observed by the eye is the same as for normal magnifying power, the numerical aperture which is actually used is reduced, and consequently, as we shall see later, the resolving power is lessened.

Problems

1. A telescope, with a single eye-lens, is 100 cm. long (objective to eyepiece). If the magnifying power is 9, find the focal length of each lens.

2. A converging lens of focal length 4 cm. is used as a magnifying glass by a person with normal eyes. What range of values can the magnifying power have without injury to the eye? What is the most desirable value of the magnifying power?

3. Given $F = 30$ cm., $f = 5$ cm., aperture of objective $= 3$ cm., aperture of eyepiece $= 2$ cm.; find

(*a*) where the eye must be placed to see the greatest extent of object;
(*b*) the diameter of the eye-ring;
(*c*) when eye is at eye-ring, the angular size of the largest object which can be seen.

4. If the aperture of eye pupil $= 3$ mm., and $F = 25$ cm., $f = 5$ cm., find the aperture of the objective so that the pupil is just completely filled with rays of light. What would be the effect of (*a*) reducing, (*b*) enlarging the aperture of the telescope?

5. When focussed on a star, the distance of the eyepiece of a telescope from the object lens is 50 cm. The focal length of the eyepiece is 1.2 cm. To see a certain tree 10 m. high the eyepiece must be drawn out 0.25 cm. What is the distance of the tree from the observer, and the size of the image formed by the object glass?

6. A distant object is observed through a telescope, with a single lens (of focal length 5 cm.) as eyepiece. If the pupil of eye is 3 mm. in diameter and it is placed at the position where the greatest length of object may be seen, find approximately the smallest aperture the telescope objective of focal length 100 cm. may have without cutting down the intensity of the image.

7. (*a*) When a *small* object placed at a point on the axis is viewed through a combination of two co-axial converging lenses 2 cm. apart, one of focal length 6 cm., the other (next the eye) of focal length 4 cm., the virtual image is 400 cm. from the lens next the eye. Find where the object is.

virtual image is 400 cm. from the lens next the eye. Find where the object is.

(*b*) If the aperture of the lens next the eye is 2 cm., find through what distance the eye could be moved along a line at right angles to and intersecting the axis at a distance of 20 cm. from the lens, and in all positions see the image.

8. A small object is viewed through a combination of two converging lenses, each of focal length 5 cm., placed 3 cm. apart. If the final image is located 30 cm. from the lens next the eye, find the *magnifying power* of the combination.

9. In a Ramsden eyepiece prove that rays will emerge "parallel" from a point placed at a distance $\frac{1}{4}f$ (where f is the focal length of each lens) from the field lens.

10. Solve 9 by making a diagram to scale. (Trace an inclined parallel beam backwards through the combination.)

11. Prove that if an object be placed at a distance $= \frac{1}{4}f$ from the field lens of a Ramsden eyepiece the magnifying power $= \dfrac{25}{\frac{3}{4}f}$.

12. What advantages has a Ramsden eyepiece over a single eye-lens?

13. If the objective of a telescope has a diameter of 4 cm., and a focal length of 30 cm., and if the focal length of each lens of the Ramsden eyepiece is 6 cm., find the diameter and the position of the eye-ring.

14. Prove that the magnifying power of a telescope equipped with a Huygens' eyepiece is $\dfrac{2F}{3f}$, where F is the focal length of the objective, and $3f$, f, are the focal lengths of the component lenses of the eyepiece.

15. When a small object is viewed through a magnifying glass of focal length 6 cm., the *angular* size of the image as viewed by an eye immediately behind the glass is $\dfrac{18}{5\pi}$ degrees. If the object is 5 cm. from the glass, find its linear size.

16. A distant object is viewed through a telescope whose objective has a focal length of 30 cm., and (single) eye-lens a focal length of 4 cm., the centres of the two lenses being separated 33.6 cm. Find, from fundamental considerations, the *magnifying power* of the telescope *for this arrangement*. Illustrate by a good diagram.

17. Under what condition is it correct to state that the magnifying power of a telescope is equal to $\dfrac{\text{diameter of objective}}{\text{diameter of exit-pupil}}$?

18. A microscope has an objective with N.A. 0.20, giving an initial magnification of 25, and an eyepiece which magnifies 5 times. How

many times greater is the actual M.P. of this microscope than its normal M.P.? (Take the diameter of the pupil of the eye as 2.5 mm.)

19. (*a*) An objective with N.A. 0.85 gives an initial magnification of 60. If this objective is used with an eyepiece of focal length 5 cm., find the radius of the exit-pupil.

(*b*) What is the normal M.P. for the above microscope?

20. A microscope has an objective of N.A. 0.50 and an eyepiece magnifying 5 times. If the initial magnification of the objective is 16, prove that the diameter of the exit-pupil exceeds that of the eye-ring. Hence show that with this arrangement the full N.A. of the objective is not utilized.

CHAPTER VI

DISPERSION

57. Refraction through a Prism. — Suppose plane waves are incident on a portion of a refracting medium in the form of a prism, the medium being more dense than the surroundings. Because of the slower velocity of the wave-disturbance in the prism, it is readily seen by applying Huygens' Principle that an

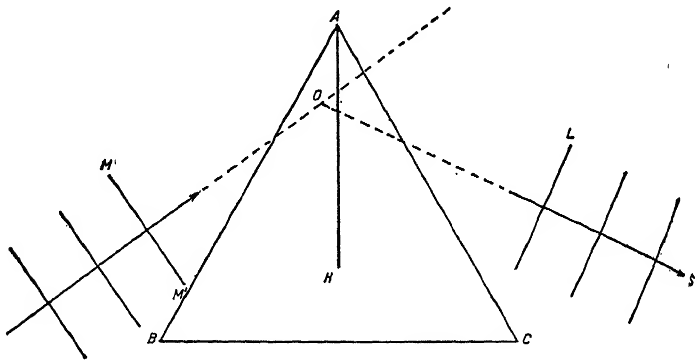


FIG. 78.

incident wave-front, such as MM', Fig. 78, gives rise, at the time the outer edge has reached A, to the wave-front AH, while, by the time the other edge has passed through the face AC, the emergent wave-front L is formed. Plane waves, therefore, incident on the prism in the direction KN emerge in the direction OS, being *deviated* through the angle NOS.

In the case of light, it is more usual to discuss such questions as the amount of deviation by dealing with rays. Suppose, then, that ray EM is one of a parallel bundle of rays incident at an angle i_1 on the face HB, Fig. 79, of a prism of index n and refract-

ing angle A, and emergent from the second face HC at an angle i_2. For the moment we are confining our attention to rays of what is called *monochromatic* light, in other words, light corresponding to a single wave-length. A good practical illustration (to a first approximation) is found in the yellow light emitted from a flame

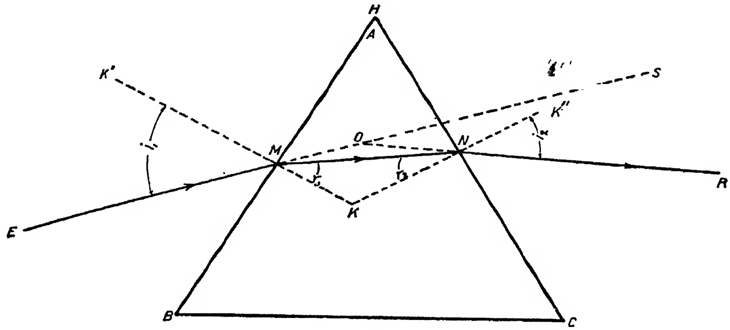


FIG. 79.

coloured with common salt. If, then, D represents the angle of deviation, we have

$$
\begin{aligned}
D &= \angle\, SOR \\
&= \angle\, OMN + \angle\, ONM \\
&= i_1 - r_1 + i_2 - r_2,
\end{aligned}
\qquad (6.01)
$$

where $K'MK$ and $K''NK$ are normals to the surfaces HB and HC and

$i_1 = \angle\, K'ME$, the angle of incidence,

$r_1 = \angle\, NMK$, the corresponding angle of refraction,

$i_2 = \angle\, K''NR$, the angle of emergence,

$r_2 = \angle\, MNK$, the corresponding angle of refraction.

Since $\qquad r_1 + r_2 + \angle\, MKN = \pi,$

and $\qquad A + \angle\, MKN = \pi,$

$$\therefore\quad r_1 + r_2 = A,$$

and we have, $\qquad D = i_1 + i_2 - A. \qquad (6.02)$

Angle of Minimum Deviation. — The magnitude of D obviously depends on the angle at which the parallel bundle of rays

is incident on the prism. If, now, when the emergent light is focussed on a screen or is observed in a telescope, the *prism* is rotated, it will be found by actual trial that there is one position of the prism and only one, for which the deviation is a minimum. When this is the case, the angles i_1 and i_2 are equal. This may be proved experimentally, or by the following *reductio ad absurdum* method.

Suppose that when the prism is in the position of least deviation, i_1 is not equal to i_2. Then, imagine the beam of light reversed in direction, that is, incident parallel to RN, emergent parallel to ME. The deviation is not altered, being still equal to $\angle SOR$, and is, therefore, still a minimum. On our hypothesis, therefore, we have minimum deviation for *two* angles of incidence i_1 and i_2.

But by experiment this cannot be the case. Therefore the hypothesis cannot be true, that is, in the position of minimum deviation

$$i_1 = i_2.$$

In much experimental work the prism is set in this position. When this is the case, frequent use is made of an important relation which we shall now derive.

Since $$i_1 = i_2 = i,$$

and $$r_1 = r_2 = \frac{A}{2},$$

$$D = 2i - A,$$

or $$i = \frac{D + A}{2}.$$

But $$n = \frac{\sin i}{\sin r},$$

$$n = \frac{\sin \dfrac{D + A}{2}}{\sin \dfrac{A}{2}} \tag{6.03}$$

Small-Angled Prisms. — If an angle is sufficiently small, we may write $\sin A = A$, in which case relations (6.02) and (6.03) reduce to

$$D = (n - 1)A. \tag{6.04}$$

The student should be careful to note that relation (6.03) is applicable accurately only to cases of minimum deviation, while relation (6.04) is restricted to prisms of small angles.

58. The Spectrometer. — Relation (6.03) provides us with one of the standard methods of measuring the index of refraction of a material, for monochromatic light, when in the form of a prism.

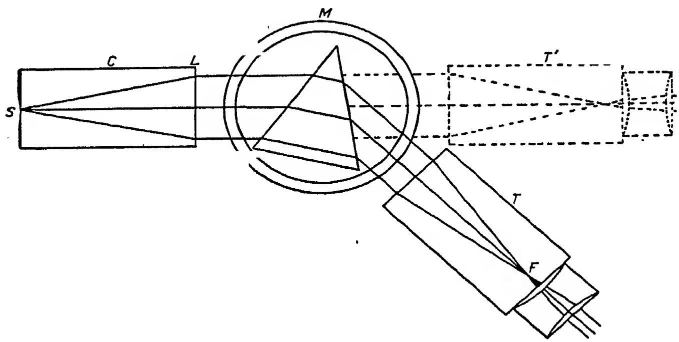


FIG. 80.

The method consists in measuring the angles A and D by means of a spectrometer. While it is not the purpose of this book to discuss details of experimental procedure, certain principles cannot be considered without a clear understanding of the fundamental ideas involved in the construction and use of this instrument.

The essential parts of a spectrometer are, (*a*) a collimator C, Fig. 80, consisting of a slit S whose width and whose distance from the converging lens L may be adjusted; (*b*) a telescope T, provided with cross-hair lines, and capable of rotation about a central vertical axis passing through (*c*), a table M, on which a prism or other objects may be placed. The position of the tele-

scope may be described by means of the circular scale whose centre is the axis of rotation. When the instrument is adjusted for ordinary use the telescope is focussed for parallel rays, while the slit S is placed so that parallel rays emerge from the collimator lens.

Fig. 80 shows the passage of a parallel beam of rays, originally emerging from a very narrow slit S, through a prism placed in the position of least deviation. It will be seen that a real image of the slit is formed at F in the focal plane of the objective of the telescope. By observing this image through the eyepiece, the cross-hairs may be set on it and the position of the telescope carefully read off the scale. If the prism is then removed, and the telescope shifted to position T' where the direct beam is received, a second scale reading can be taken. The difference between the two readings is, of course, equal to D. Since the angle A can readily be found (using the same instrument in a manner not necessary to describe in this book) the index n can then be evaluated.

Values of indices of refraction, for a few materials, for yellow (D) light will be found in Table IV.

Table IV

Refractive Index of Various Glasses for D Light

Glass	Trade Number	Density	Index
Ordinary Crown....	0.203	2.54	1.5174
Borosilicate Crown..	0.144	2.47	1.510
Very Heavy Crown..	0.2994	3.60	1.613
Extra Light Flint...	0.726	2.87	1.540
Ordinary Flint......	0.118	3.58	1.613
Very Heavy Flint...	0.198	4.99	1.778
Extra Heavy Flint..	S.386	6.01	1.917

Should the slit S be made wider, the student should realize that from each point in the plane of the slit a separate bundle

PLATE II

FIG. 1. Three spectra of magnesium, with different degrees of excitation (Electronic bombardment). (Foote, Meggers and Mohler.)

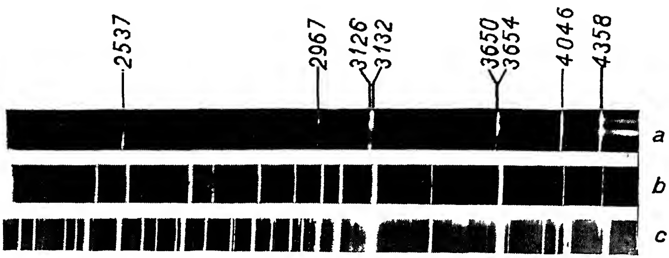


FIG. 2. Three spectra of mercury with different degrees of excitation.

FIG. 3. A typical unresolved band spectrum.

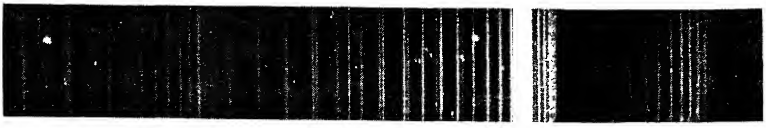


FIG. 4. To show individual lines of a single band. (Pearse.)

will spread out, each giving rise to its own parallel beam emerging from the collimator lens. It follows, therefore (as is represented in Fig. 81), that for each point in the slit, such as S_1 and S_2, there are corresponding point images in the focal plane of the objective,

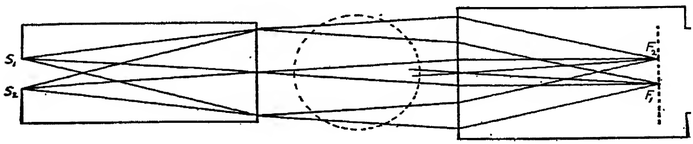


FIG. 81.

such as F_1 and F_2. In other words the wider the slit used, the wider is the image observed in the telescope. (Try problem 15 at end of chapter.)

59. Dispersion. — Suppose, now, instead of using as our source of light a flame coloured with common salt and thus giving us a yellow slit image, we change to a flame coloured by another salt, lithium chloride, for example. We then observe a *red* image of the slit and one, moreover, in a position not the same as the yellow. On measuring the angle of minimum deviation for this light, we find a value less than that obtained for yellow light. As a result of such observations, we reach the important conclusion that the velocity of red light *in the prism* is more than the velocity of yellow.

If once more we change our source of light, this time to a mercury arc lamp, we now see two yellow slit images, a very bright green, a bright violet, and possibly a few others, all of course in different positions. We conclude that the mercury arc emits several different kinds of light, each of which travels in the prism with its own velocity. Moreover, by actual measurement of the value of n for each slit image, we can readily find the corresponding velocity. This collection of slit images for any given source of light constitutes its spectrum. In Plate I and Plate II the student will find actual photographs of various spectra.

White Light. — If the slit is illuminated with sunlight (see, however, section 72) or by an incandescent lamp, one sees the familiar "rainbow" spectrum with colours ranging continuously from red to violet. This classic experiment, in a slightly different manner, was first performed by Sir Isaac Newton, who drew the important conclusion that "light is not similar or homogeneal but consists of difform rays, some of which are more refrangible than others. . . ." It is perhaps not out of place to state that in the ante-chapel of Trinity College Chapel, Cambridge, England, a life-size statue of Newton represents him holding a prism in his hand.

What now, we ask, physically distinguishes one colour from another, or even a slit image in one colour from another of the same colour but in a slightly different position? If we consider light as a wave phenomenon, the answer is not difficult. Since it is an experimental fact that *in free space* light of all colours travels at the same rate, it follows that differences in colour must be due to differences in frequency and in wave-length, one of which is inversely proportional to the other (at constant velocity). We describe a kind of light, therefore, by stating either the frequency giving rise to it, or the resulting wave-length in some standard medium. Light of a particular quality then corresponds to a note in sound. The latter is usually described by giving its frequency but might be described by giving its wave-length in air. In light it is usual to describe the quality in terms of wave-length in air at 15° C. and 760 mm. pressure, although sometimes frequency is used.

For example, the two yellow slit images in the case of mercury correspond to air wave-lengths 0.00005790 cm. and 0.00005770 cm.; the bright green to 0.00005461 cm.; a violet to 0.00004358 cm. Methods of obtaining such values will later be discussed in detail. At this stage it may be pointed out that certain important wave-lengths, of which we shall make some use in this chapter, have been designated by letters. A few of these, as well as the luminous substance giving rise to them, are tabulated for reference in Table V.

The Angstrom. — Because of the shortness of light waves the unit of length generally used is the angstrom or 10^{-8} cm. The

wave-length of the mercury green line, for example, is usually given as 5461 angstroms or 5461 A, or sometimes just 5461. The name angstrom is in honour of Ångstrom (Sweden, 1814–1874), a man who made a chart of the solar spectrum which for a number of years was used as a standard in the measurement of wave-lengths. The unit of wave-length adopted by international agreement is called the *international angstrom* (I.A.). For ordinary

Table V

Designation of Certain Wave-Lengths

Designation	Colour	Source	Wave-length
A'	red	potassium	7682 A
A	red	oxygen	7594
C	red	hydrogen	6563
D_1	yellow	sodium	5896
D_2	yellow	sodium	5890
F	blue	hydrogen	4861
G'	violet	hydrogen	4340
G	violet	calcium	4308
H	ultra-violet	calcium	3968

purposes, there is no objection to defining this in the way we have just done, that is, as 10^{-8} cm., but more accurately it is defined in terms of the wave-length of a red spectral line, emitted by cadmium, in dry air, at 15° C., 760 mm. pressure, with $g = 980.67$. This particular wave-length has been measured with the highest degree of accuracy, and the value of 6438.4696 I.A. was adopted as a standard by the International Astronomical Union.

The student must not fail to realize that because of the change in velocity when we pass from one medium to another, the wave-length of any given kind of light will also change. In water, for example, the wave-length of the green mercury line is only approximately three-quarters of 5461 A. Even in air wave-lengths are not exactly the same as in free space, although for some pur-

poses the difference is so slight that we may neglect it. This change in wave-length with changing medium, for a given quality of light, must not be confused with the dispersion of a medium. *The dispersion of a medium refers to the change in velocity in that medium when we change the kind of light, or, as we may now more accurately state it, when we change the wave-length.* Because of this change in velocity, we have the difference in deviation already noted. Indeed, sometimes dispersion is thought of as the change

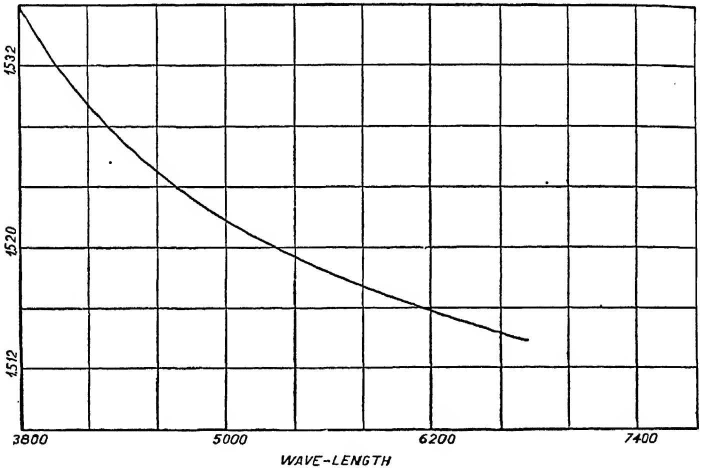


FIG. 82.

in deviation resulting from changing wave-length. If a prism is set in a fixed position, the deviation measured for a series of wave-lengths, and a graph made by plotting deviation against dispersion, we obtain what may loosely be called a dispersion curve for that prism. More accurately, however, a dispersion curve for any substance is obtained by plotting the index of refraction against the wave-length. Fig. 82 is a graph obtained by making use of data for the crown glass given in Table VI.

As an exact measure of the dispersion of a medium, we use the value of $\frac{dn}{d\lambda}$, which, it will be seen, is just a measure of the

Table VI

Refractive Indices for Certain Materials

(from Kaye and Laby tables)

Wave-Length	Light Jena Crown Glass (15° C.)	Dense Jena Flint Glass (15° C.)	Rock Salt (18° C.)	Water (20° C.)
6708 A	1.5140	1.6434	1.5400	1.3308
6563	1.5145	1.6444	1.5407	1.3311
5893	1.5170	1.6499	1.5443	1.3330
5461	1.5191	1.6546	1.5475	1.3345
4861	1.5230	1.6637	1.5534	1.3371
4047	1.5318	1.6852	1.5665	1.3428

slope of the graph at any given point. The student should also note that the dispersion is greater in the region of the shorter wave-lengths at the violet end of the spectrum than at the red end. (Examine again Fig. 82.)

Dispersion Formulae. — The graph of Fig. 82 illustrates a case of *normal* dispersion as distinguished from anomalous dispersion, a question which will be considered in a later chapter. This graph may be represented mathematically by simple empirical relations, two of which will be noted. According to the first, due to Cauchy, we write

$$n = A + \frac{B}{\lambda^2},^* \tag{6.05}$$

where A and B are constants for any given medium. For example, for a certain kind of glass, $A = 1.616$, $B = 0.100 \times 10^{-9}$, and we may therefore, to a fair degree of accuracy, at once calculate the value of n for any wave-length.

* More accurately $n = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \cdots$ but for many purposes relation (6.05) is sufficiently accurate.

By simple differentiation, we have from relation (6.05)

$$\frac{dn}{d\lambda} = -\frac{2B}{\lambda^3}. \tag{6.06}$$

In the case of the glass whose constants have been given, for G light, $\frac{dn}{d\lambda} = -2500$. (What is the significance of the minus sign?)

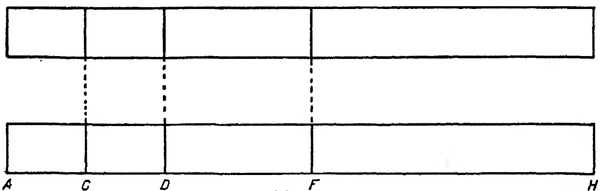
A second dispersion formula, which is useful if not applied to too great a range of wave-lengths, is the following due to Hartmann:

$$n = n_0 + \frac{c}{\lambda - \lambda_0}. \tag{6.07}$$

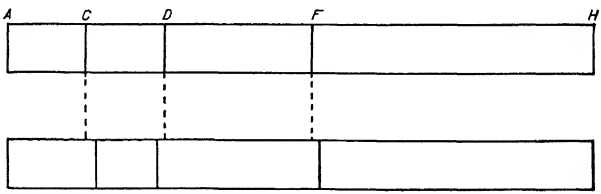
To take another concrete example, for a certain glass,

$$n_0 = 1.58882; \; \lambda_0 = 2.2906 \times 10^{-5}; \; c = 0.10113 \times 10^{-5}.$$

60. Irrational Dispersion. — Suppose prisms of two different substances are chosen with refracting angles such that the angular separation between spectral lines A and H is exactly the same for



(a)



(b)

FIG. 83.

each. It might then be expected that the angular separation between A and some intermediate wave-length such as F would also be the same in the spectrum produced by each prism — as represented in *a*, Fig. 83. As a matter of fact this is not the case, the actual state of affairs being somewhat as represented in *b*, Fig. 83. It will be noted that intermediate spectral lines do not occupy exactly corresponding positions, that is, one spectrum is not an exact duplicate of the other. This feature of prismatic spectra is what is known as *irrational dispersion*. While the magnitude of the effect is not great, it is important and the student should note carefully the numerical example given in Table VII. Here the numbers give the angular separations in minutes between A light and each of four other wave-lengths, for two different kinds of glass prisms so chosen that the angular width between A and H lines is the same for each.

Table VII

Angular Width between A Line and each of C, D, F, and H, for Two Prisms

	A	C	D	F	H
Crown glass........	0	1.8	3.6	7.2	13.8
Flint glass.........	0	1.9	3.45	7.3	13.8

61. Spectrometer, Spectroscope, Spectrograph. — From the brief description of the essential features of a spectrometer given above, it should not be difficult to see that, in addition to its use for the measurement of refractive indices, it is also of use in observing a spectrum, that is, the various wave-lengths emitted by a source of light. An instrument specially designed for such visual observation is called a *spectroscope*.

It is frequently desirable to have a permanent record of the spectrum of a source, in which case a *spectrograph* is used. In a prism spectrograph essentially all that is done is to remove the eyepiece of the telescope and to place a photographic plate in

the focal plane of the objective lens. Fig. 84 shows the arrangement of such an instrument, while the following problem should make clear some of the fundamental ideas better than detailed description.

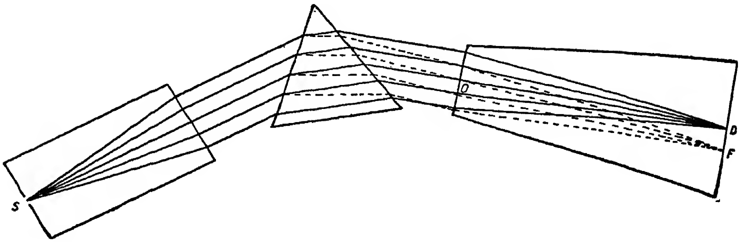


FIG. 84.

A 60° flint glass prism, having an index = 1.6222 for D light, and = 1.6320 for F light, is set in the position of minimum deviation for D light. (a) When the incident light consists of a beam of parallel rays, find the angular separation between the emergent D and F beams. (b) If the emergent light is focussed on a screen (or camera plate) by an achromatic lens of focal length 60 cm., find the linear distance (or length of spectrum) between the D and the F images.

Part (*a*).

Since the deviation is a minimum for D light, therefore, for this light,

$$r_1 = r_2 = 30^\circ,$$

$$\frac{\sin i_1}{\sin 30^\circ} = 1.6222,$$

from which $\qquad i_1 = 54^\circ\ 12'.$

$\therefore$ for D light $\qquad i_2 = 54^\circ\ 12'.$

To find i_2 for F light, it must be remembered that for this wave-length, the prism is not in the position of least deviation (although very approximately so). We must, therefore, deal with refraction at each surface separately.

At the first surface, since i_1 is the same for all wave-lengths,

$$i_1 = 54° 12'$$

$$\frac{\sin 54° 12'}{\sin r_1} = 1.6320,$$

from which
$$r_1 = 29° 48'.$$

But
$$r_1 + r_2 = 60°,$$
$$r_2 = 30° 12'.$$

Finally, since
$$\frac{\sin i_2}{\sin 30° 12'} = 1.6320,$$
$$i_2 = 55° 11'.$$

Therefore, the required angular separation between the F and D images

$$= 55° 11' - 54° 12' = 59'.$$

Part (*b*).

Referring to Fig. 84, we see at once that if D represent the D image of the slit, F the F image,

$$\angle FOD = 59'.$$

Since this is a small angle, we can put $OF = OD = 60$ cm., and write

$$FD = 60 \tan 59'$$

or
$$= 60 \sin 59'$$

or
$$= 60 \cdot (\text{angle } FOD \text{ in radians})$$
$$= 1.03 \text{ cm}.$$

62. Dispersion of an Instrument. — Sometimes the word dispersion has reference to the separation of any two wave-lengths produced by an instrument, and in that case is measured by the magnitude of $\frac{d\theta}{d\lambda}$.

For example, if θ_1 = deviation of wave-length λ_1,
and θ_2 = deviation of wave-length λ_2,

then the quantity $\frac{\theta_2 - \theta_1}{\lambda_2 - \lambda_1}$ is a measure of the average dispersion in the region λ_1 to λ_2, provided the wave-lengths are fairly close together. In the limit, when λ_2 differs from λ_1 by a small quantity, this expression becomes $\frac{d\theta}{d\lambda}$. Its actual magnitude, which may be expressed in such units as radians per centimetre, degrees per centimetre, radians per angstrom, obviously depends on, (*a*) the refracting angle of the prism, (*b*) the material of which the prism is made, that is, the dispersion of the medium. As the solution of the above problem shows, the linear separation between any two images also depends on the focal length of the camera lens.

Sometimes the dispersion of an instrument is also expressed in terms of the number of angstroms corresponding to a distance of 1 mm. on the photographic plate. In the above problem, for example, we can state that in the region between the F and D lines, the average dispersion is

$$\frac{5893 - 4861}{10.3}, \text{ or } 100 \text{ angstroms per mm.}$$

63. Achromatism. — From the general fact of dispersion, as well as from the standard lens relation $\frac{1}{f} = (n - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$, it follows that the focal length of a lens varies with the wave-length. For example, if we are dealing with a double convex

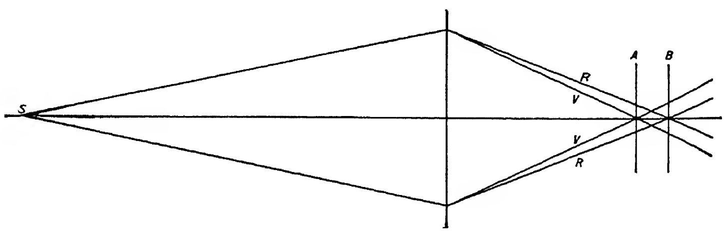


FIG. 85.

crown glass with each face of 20 cm. radius, and with indices as given in Table VI, the student can readily verify that the focal length for wave-length 6708 A is 19.46 cm., while for wave-

length 4047 A it is 18.80 cm. This variation of the focal length because of dispersion is called *chromatic aberration*. Its existence may be nicely shown by the arrangement of Fig. 85, where light from a small white source S falls on a lens of tolerably wide aperture and is received on a ground glass screen, first in position A, then in position B. In the former case the outer edge of the spot on the ground glass is reddish, in the latter, blue or violet.

Because of this defect it is obviously impossible to obtain perfect images with any instrument built of simple spherical lenses. Coloured edges appear where there is no colour in the original source, and perfect focussing cannot be obtained. The very important question then arises, Is it possible to correct for chromatic aberration? In other words, Can we make an *achromatic* lens, one whose focal length is the same for all colours? Because of differences in *dispersive power* of various kinds of glass, this can readily be done, and it is necessary to consider this question somewhat in detail.

If two prisms, each of the same refracting angle, one of crown glass, the other of flint, are used to project the spectrum of a source of white light on a screen, two things are observed. (*a*) As might be expected, because of the higher refractive index, the average deviation is greater in the case of the flint than of the crown glass. (*b*) The total length of spectrum is considerably greater for the flint than for the crown glass prism. To make the matter concrete, consider two 10° prisms with indices as given in Table VI. Then, using the simple relation $D = (n - 1)A$, we have, for crown glass,

$$D, \text{ for } \lambda 6708 = (1.5140 - 1)10° = 5.14°,$$
$$D, \text{ for } \lambda 4047 = (1.5318 - 1)10° = 5.318°.$$

Therefore, the angular width between these two wave-lengths $= 0.178°$.

Similarly, for flint glass,

$$D, \text{ for } \lambda 6708 = (1.6434 - 1)10° = 6.434°,$$
$$D, \text{ for } \lambda 4047 = (1.6852 - 1)10° = 6.852°,$$

and angular width between the same two wave-lengths $= 0.418°$.

We see, then, that the length of the spectrum due to the flint glass prism is over twice as great as that due to the crown, although the deviation of any one colour is only slightly greater in the case of the flint. Flint glass, therefore, is said to have a greater dispersive power than crown glass. By comparing different transparent substances in this way, marked differences are found in dispersive powers.

Because of this fact, it is possible to choose a combination of two different kinds of prisms so that one counteracts the dispersion of the other, while a resultant deviation remains. We then have an *achromatic combination* of two prisms, giving deviation without dispersion. Again, a concrete numerical example will make the point clearer.

Problem. *Using the constants given in Table VI, find the angle of a flint glass prism which gives the same length of spectrum between C and F wave-lengths as a 10° crown glass prism.*

Since $D_C = (n_C - 1)A$, where the subscript indicates the wave-length, we have for crown glass

$$\begin{aligned} D_F - D_C &= (n_F - n_C)A \\ &= (1.5230 - 1.5145)10^\circ \\ &= 0.085^\circ. \end{aligned}$$

Also for flint glass,

$$\begin{aligned} D_F - D_C &= (1.6637 - 1.6444)A \\ &= 0.0193^\circ A. \\ \therefore \quad 0.0193^\circ A &= 0.085 \end{aligned}$$

or

$$A = 4.40^\circ.$$

If, then, these two prisms are placed together as in Fig. 86, it follows that one prism will counteract the dispersion of the

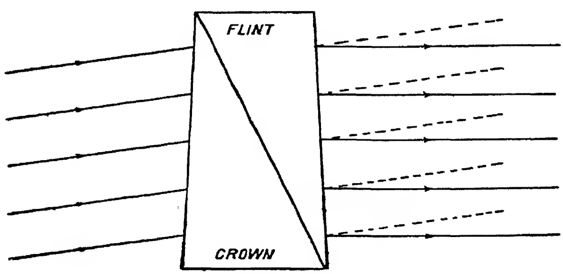


FIG. 86.

other, while, because of the much smaller refracting angle of the higher dispersive flint prism, there will be a resultant deviation in the same direction as that due to the crown.

To find the resultant deviation, we have at once,

$$\text{net deviation of } C \text{ light} = (1.5145 - 1)10 - (1.6444 - 1)4.40 = 2.31^\circ$$

$$\text{net deviation of } F \text{ light} = (1.5230 - 1)10 - (1.6637 - 1)4.40 = 2.31^\circ$$

$$\text{net deviation of } D \text{ light} = (1.5170 - 1)10 - (1.6499 - 1)4.40 = 2.31^\circ$$

In making achromatic combinations for visual purposes usually C and F wave-lengths are used. To a first approximation, it does not matter what two wave-lengths are used, but because of irrational dispersion, slight differences in the refracting angles may be found for different pairs of wave-lengths. Irrational dispersion is also responsible for the fact that a combination of two prisms may not be *perfectly* achromatic for every wave-length in the spectrum. The achromatism may be made more perfect by using three different prisms and three wave-lengths.

64. Achromatic Lens. — The same principle is applied to lenses. Thus, if a converging crown lens is combined with a diverging flint of higher dispersive power, dispersion of the first lens can be counteracted by the second and a resultant converging lens obtained with focal length the same for all colours.

In dealing with this question quantitatively it is well to note the following:

(*a*) When two thin lenses are in contact, if
f_1 = focal length of the first, for any wave-length,
f_2 = focal length of the second, for the same wave-length,
then F the resultant focal length is given by

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}. \tag{6.08}$$

(*b*) The condition for achromatism, using C and F wave-lengths, gives at once

$$F_C = F_F \tag{6.09}$$

Bearing these relations in mind, as well as the standard lens relation

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right),$$

any ordinary achromatic lens problem may readily be solved. One example will be given.

Problem. *A converging crown glass lens and a diverging flint form a converging achromatic lens of focal length 50 cm. Using the indices given in Table VI, find the focal length of each lens for C, D, and F light.*

We are given

$$F_C = F_F = +50,$$

$$\therefore \quad +\frac{1}{50} = \frac{1}{f_1} + \frac{1}{f_2} \text{ for any wave-length.}$$

Therefore, using C light,

$$+\frac{1}{50} = (1.5145 - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + (1.6444 - 1)\left(\frac{1}{r_3} - \frac{1}{r_4}\right),$$

and using F light

$$+\frac{1}{50} = (1.5230 - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + (1.6637 - 1)\left(\frac{1}{r_3} - \frac{1}{r_4}\right),$$

where r_1, r_2 are radii of curvature of faces of crown lens and r_3, r_4 are radii of curvature of faces of flint lens.

Putting $\quad \left(\frac{1}{r_1} - \frac{1}{r_2}\right) = x, \text{ and } \left(\frac{1}{r_3} - \frac{1}{r_4}\right) = y,$

we have the two equations

$$+\frac{1}{50} = 0.5145x + 0.6444y,$$

$$+\frac{1}{50} = 0.5230x + 0.6637y,$$

from which $\quad x = +0.08773, \; y = -0.03863.$

Finally, for the crown glass lens, we have

$$\frac{1}{f_C} = 0.5145 \times 0.08773, \text{ or } f_C = +22.16 \text{ cm.}$$

$$\frac{1}{f_D} = 0.5170 \times 0.08773, \text{ or } f_D = +22.05 \text{ cm.}$$

$$\frac{1}{f_F} = 0.5230 \times 0.08773, \text{ or } f_F = +21.79 \text{ cm.},$$

while for the flint lens

$$\frac{1}{f_C} \quad 0.6444 \times (-0.03863) \text{ or } f_C = -40.2 \text{ cm.}$$

$$\frac{1}{f_D} = 0.6499 \times (-0.03863) \text{ or } f_D = -39.8 \text{ cm.}$$

$$\frac{1}{f_F} = 0.6637 \times (-0.03863) \text{ or } f_F = -39.0 \text{ cm.}$$

65. Measure of Dispersive Power. — Achromatic lens problems are sometimes solved by making use of a numerical measure of dispersive power. What such a measure is obviously must be a matter of definition. Since, in the case of a prism, the actual length of a spectrum depends on the material of which it is made, and moreover increases as the refracting angle is increased (or as the mean deviation is increased), it is agreed to measure the dispersive power, usually represented by ω, by the spectrum length between C and F wave-lengths, per unit angle of mean deviation. In symbols,

$$\omega \quad \frac{D_F - D_C}{D_r} \tag{6.10}$$

the mean deviation being taken as that of D light, intermediate between C and F.

By using the relation giving the deviation for small-angled prisms, this may be written

$$\omega = \frac{(n_F - 1)A - (n_C - 1)A}{(n_D - 1)A}.$$

$$\frac{n_F - n_C}{n_D - 1} \tag{6.11}$$

It will be seen that this expression, involving only indices, is a function of the medium alone, and, therefore, is a suitable general measure of dispersive power.

66. General Condition for an Achromatic Combination of Two Lenses. — Making use of notation which now should need no explanation, we have, for C light, and two lenses in contact,

$$\frac{1}{F_C} = (n_{1C} - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + (n_{2C} - 1)\left(\frac{1}{r_3} - \frac{1}{r_4}\right).$$

Also for F light:

$$\frac{1}{F_F} = (n_{1F} - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + (n_{2F} - 1)\left(\frac{1}{r_3} - \frac{1}{r_4}\right).$$

Since $F_C = F_F$,

$$(n_{1F} - n_{1C})\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + (n_{2F} - n_{2C})\left(\frac{1}{r_3} - \frac{1}{r_4}\right) = 0.$$

$$\frac{(n_{1F} - n_{1C})}{(n_{1D} - 1)}(n_{1D} - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + \frac{(n_{2F} - n_{2C})}{(n_{2D} - 1)}(n_{2D} - 1)\left(\frac{1}{r_3} - \frac{1}{r_4}\right) = 0.$$

$$\therefore \quad \frac{\omega_1}{f_1} + \frac{\omega_2}{f_2} = 0, \tag{6.12}$$

where f_1 and f_2 represent the mean focal lengths, that is, for wave-length D, of the lenses.

Problems dealing with achromatic lenses are frequently solved using this general condition, but, in the writer's opinion, it is preferable to use the more fundamental method illustrated in the solution of the above problem.

67. Direct Vision Spectroscope. — It is sometimes convenient to have a spectroscope which gives a spectrum with but little deviation, — indeed for one wave-length, with no deviation at all. This can be done by a proper choice of two or more (generally more) prisms. The problem is a very simple one, for, taking the

case of two prisms as an example, all that is necessary is to choose the refracting angles so that the resultant deviation is zero for some wave-length about the middle of the spectrum. Red light is then deviated a little to one side of the direction of the incident light, violet to the other. A numerical example will be found in problem 18 at the end of this chapter.

Problems

NOTE. — Unless otherwise stated, use the following indices in the solution of the problems in this exercise.

	A	*C*	*D*	*F*	*H*
Crown glass......	1.528	1.531	1.534	1.540	1.551
Flint glass.......	1.578	1.583	1.587	1.597	1.614

1. A source emits waves of length 6563 A. What is its frequency?

2. If a 60° prism is set in the position of minimum deviation for F light (for which $n = 1.534$), find the angle at which light leaving the collimator tube of the spectrometer strikes the first face of the prism.

3. Light from a source containing lithium and sodium illuminates the slit of a spectrometer. When a 60° prism is set successively in the position of minimum deviation for the red lithium line and the D sodium, an angular difference in the values of the minimum deviation of 1° 30′ is obtained. If the index of refraction of the prism for the D line is 1.612, find the index for the red line.

4. Two spectral lines λ_1 and λ_2, $\lambda_2 < \lambda_1$, are separated a distance equal to 1 mm. on a spectrogram taken with a spectrograph whose camera lens has a focal length of 50 cm. If the index of refraction for λ_1 is 1.620 and a 60° prism is used, find the index for λ_2. (Assume minimum deviation for λ_1)

5. A 60° prism has an index of refraction of 1.622 for the D spectral line and of 1.635 for the F line. If white light is incident at an angle of 45°, what are the respective angles of emergence for these two colours?

6. When a 60° prism is set in the position of least deviation for C light, the angular separation of C and F images is 2°. If the index of refraction for C light is 1.576, indicate how you would find the index for F light.

7. (*a*) Using the same indices as in problem 5, find the minimum deviation for each colour.

(*b*) If the spectrometer telescope has a focal length of 30 cm. what is the length of the spectrum between the D and F lines when the prism is set for minimum deviation for the D lines?

8. Parallel rays of white light fall on an 8° flint glass prism, and the emergent light is focussed on a screen by an achromatic lens of focal

length 30 cm. Assuming minimum deviation, calculate the length of the spectrum between the *A* and the *H* lines.

9. Parallel rays of white light are incident at an angle of 45° on a glass prism with an angle of 60°. If the index of refraction of the *C* line is 1.66 and that of the *F* line is 1.68, find the width of the spectrum between the *C* and *F* lines when the light is focussed on a screen by a lens of 100 cm. focal length.

10. The narrow slit in the collimator of a single prism spectrograph is illuminated with white light and the prism is set in position of minimum deviation for *C* spectral line. Make use of the following data to calculate, (*a*) the deviation of the *C* line, (*b*) the length of the spectrum between the *C* and *H* lines.

Angle of prism = 60°, index for C = 1.583, index for H = 1.614, focal length of focussing lens of spectrograph = 30 cm.

11. Make a complete diagram of the optical system of a spectrograph. Trace the path of bundles of rays corresponding to two different wave-lengths, and use your diagram to show how you could calculate the linear length of the spectrum included between these two wave-lengths.

12. White light passes from a wide slit and then through a lens from which the rays emerge parallel. If this light falls on a prism and the refracted light is then focussed on a screen by means of a second lens, a white image with only the edges coloured is formed on the screen instead of a spectrum. Explain this with the aid of a sketch.

13. If the slit in problem 12 were narrow, and the *second* lens were removed what would be the result?

14. A parallel beam containing light of two wave-lengths λ_1 and λ_2 falls on a 60° prism placed in the position of minimum deviation for λ_1. When the emergent light is focussed on a screen by a lens of focal length 30 cm. the linear distance between the two images is 2.61 cm. Find the total deviation for λ_2 ($\lambda_2 < \lambda_1$).

Index for λ_1 = 1.60.

15. A spectrograph is provided with two narrow slits 1 mm. apart. If the focal length of the collimating lens is 30 cm., that of the camera lens 40 cm., find the distance apart of the two images in the focal plane of the camera lens.

16. Given a dense flint glass prism of angle 5°, find approximately the angle of a crown glass prism to give deviation without dispersion. Find also the deviation. Use *C* and *F* light.

17. (*a*) What is the angle of a prism of heavy flint glass that will achromatize an 8° crown glass prism for the *C* and *F* lines?

(*b*) Calculate the angles of deviation for the five rays *A*, *C*, *D*, *F*, *H*. Account for any differences you may find.

(*c*) Achromatize the same combination for *A* and *H* lines. Account for the difference in the result, if any.

18. (*a*) A beam of parallel rays of white light falls on a set of two prisms, one of crown angle = 10°, the other of flint glass, arranged so as to give no deviation for the D spectral line. What is the angular width of the spectrum between the C and F lines?

(*b*) If the emergent light is focussed on a screen by a lens of 60 cm. focal length, what is the linear width between these lines?

19. A small-angle crown glass prism forms an achromatic combination (C and F light) with a flint glass prism, the resultant deviation being 2°. Find the refracting angles of the two prisms.

20. A converging and a diverging lens, of different materials, form an achromatic combination. If the combination is to be converging, which lens must have the greater dispersive power? Explain.

21. (*a*) Find the difference in the focal lengths, for C and F wavelengths, of a double crown glass lens, each of whose faces has a radius of curvature equal to 20 cm.

(*b*) Find the focal length of an achromatic combination formed by combining this lens, in the usual way, with a concave flint lens.

22. Given a crown glass converging lens of mean focal length 20 cm., find the mean focal length of a flint glass lens to give an achromatic combination for C and F light. Find also the focal length of the combination.

23. A converging crown lens and a diverging flint lens, having a common surface of radius of curvature 25 cm., form an achromatic converging combination (C and F light) of focal length 50 cm. Find the radius of curvature of each of the other faces of the two lenses.

24. A crown and a flint glass lens form an achromatic converging combination of focal length 35 cm. (C and F light). Find (*a*) the mean (D light) focal length for each, (*b*) the focal lengths of each, for C, F, and H light.

25. A crown glass converging lens and a flint glass lens form an achromatic converging lens. If, for C light, the focal length of the crown lens is 24 cm., find, (*a*) the focal length of the combination, (*b*) the focal length of each lens for F light.

26. An achromatic combination is made of a converging crown glass lens whose focal length for D light is 20 cm., and a diverging flint whose focal length for D light is 60 cm. Find the index of the crown glass for C light.

n_D for crown = 1.559; n_D for flint = 1.637.

n_C for flint = 1.632.

27. It is desired to make an achromatic lens (for C and F light), of focal length 100 cm., by means of a convex crown glass lens combined with a plano-concave flint, the two surfaces in contact having the same radius of curvature. Calculate the necessary radii of curvature of the surfaces of the crown glass lens (combination is converging).

28. When a double convex crown glass converging lens, whose faces have the same curvature, is placed in contact with a diverging lens of mean focal length 40 cm., the mean focal length of the (converging) combination is 40 cm. If the mean index of the crown lens is 1.500, find the radius of curvature of each of its surfaces.

29. A beam of parallel rays of monochromatic light passes successively through two prisms, with refracting angles 60° and 30°, for each of which the index of refraction is 1.624. If each prism is placed so that the deviation is a minimum, show how they must be placed with respect to each other and to the beam.

30. Parallel rays of white light fall on a double-convex converging lens, of aperture 2 cm., whose index for red light is 1.540 and for violet light 1.567. If a screen is placed normal to the emergent light at a distance from the lens equal to its focal length for violet light, find the radius of the patch of light received on the screen. Take radii of curvature of the lens faces to be 20 cm. and 30 cm., and neglect spherical aberration.

31. In the case of a small-angled prism, prove that the rate at which the deviation varies with the wave-length is equal to the angle of the prism multiplied by the rate at which the index of refraction of the prism varies with the wave-length.

32. Prove that, whether the deviation is a minimum or not, provided all angles involved are small,

deviation $= (n - 1)A$, where A is the angle of the prism.

33. (*a*) Parallel rays of light, which is a mixture of λ_1 and λ_2, emerge from the collimator of a spectrograph, pass through a prism of refracting angle 10°, placed approximately in the position of minimum deviation, and the emergent light is photographed by a camera whose lens has a focal length of 60 cm. It is found that the images of λ_1 and λ_2 are $\frac{1}{2}$ mm. apart. If the index for λ_1 is 1.623, find the index for λ_2. $(\lambda_2 < \lambda_1)$

(*b*) Illustrate the above problem by a carefully drawn diagram.

34. A spectrometer whose collimator lens (adjusted in the standard way) has a focal length of 30 cm. is equipped with two very narrow slits illuminated with monochromatic light, whose centres are 2.619 mm. apart. A prism, whose index of refraction for this kind of light is 1.5321, is set in the position of least deviation *for one of the two beams*, and the emergent light is photographed in the focal plane of a lens of focal length 50 cm. Find (i) the linear separation of the two images, (ii) the error, if this separation is calculated on the assumption that the prism is in minimum deviation for both beams.

CHAPTER VII

SOME FACTS CONCERNING THE SPECTRUM

While the prism spectrograph (or spectroscope) is by no means the only instrument used for examining the wave-lengths emitted by a luminous source, at this stage a knowledge of other instruments is not necessary for the consideration of some outstanding facts revealed by such an examination. In this chapter, therefore, a brief outline is given of some of the fundamental facts of spectroscopy, their theoretical significance being postponed until our general subject has been further developed.

68. Light Sources. — It is first of all desirable for the student to be familiar with methods frequently used for making a substance luminous. The more important of these are the following.

(*a*) **The Incandescent Solid.** — Any solid substance, if sufficiently heated, first begins to glow with a reddish colour, and eventually, if the temperature goes high enough, becomes "white hot." An incandescent filament lamp is a familiar example of such a light source. The spectrum of such a source is continuous, that is, does not consist of a number of distinct spectral lines.

(*b*) **The Gas Burner.** — By introducing many substances into the flame of a non-luminous Bunsen or other gas burner, luminosity characteristic of the substance results. The spectrum of such a source is called the *flame spectrum* of the substance. It consists of only a few spectral lines, sometimes only a single line.

(*c*) **The Arc.** — The familiar electric arc between two carbon rods illustrates the basis of another common light source. A photograph of the luminous vapour between the rods gives the carbon arc spectrum, while, if some volatile metal or metallic salt is introduced into the arc, the resulting spectrum shows arc spectral lines of such a metal. If the carbon rods are replaced by iron ones, an iron arc results, while if a special arrangement is used

so that an electric current can pass through mercury vapour in an evacuated vessel, we have a mercury vacuum arc. In general there are various ways in which arc spectra may be obtained but all are characterized by currents of several amperes, with a voltage drop (potential difference) probably considerably less than 100 volts between the terminals.

(*d*) **The Spark.** — If two pieces of metal such as MM', Fig. 87, are joined to the terminals of an induction coil or small transformer, the spark which "jumps" between them on discharge is a source of light, which, when analyzed gives the *spark spectrum*

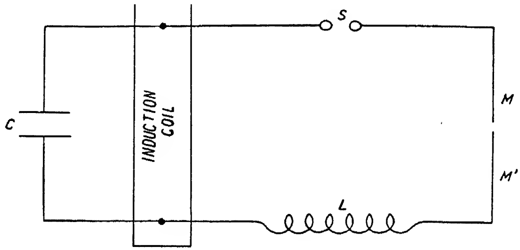


FIG. 87.

of the metal. The intensity of the spark is greatly increased by placing a condenser C in parallel with the coil, while it may be further altered by varying, (1) the length of an auxiliary spark gap S, (2) the amount of inductance L in the circuit. In the case of a spark, electrical conditions are very different from those obtaining in the arc. In the spark we have to do with small currents, fractions of an ampere, but with voltages which may be as high as many thousands. In Fig. 3, Plate I, *b* is a small portion of the arc spectrum of the metal tin, taken with a prism spectrograph, while *a* is the spark spectrum for the same region taken on the same photographic plate in such a way that it overlaps, to a considerable extent, the arc spectrum. Reference will again be made to this figure.

(*e*) **The Vacuum Tube.** — This is a very common source for the analysis of the light emitted by luminous gases and vapours. If a tube with electrodes E, E', Fig. 88, containing a gas at low

pressure (of the order of 1 mm.) is joined directly to an induction coil or transformer, a brilliant luminosity results. Tubes of various shapes may be used, but in all the principle is the same. In Fig. 1, Plate I, a reproduction of the vacuum tube spectrum (visible) of hydrogen is given, while Fig. 2, Plate I, is a similar spectrum for helium.

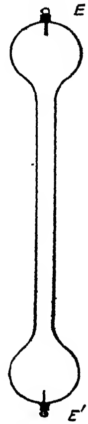


FIG. 88.

(*f*) **Electronic Bombardment.**—To explain this method, which, while not so frequently used as the others, has yielded very valuable spectroscopic information, it is necessary first of all to digress to refer to a conception of an atom which has proved extremely useful in explaining much of the experimental work of the last generation in all branches of physics. According to it, in an atom we have a planetary system on a microcosmic scale. The *nucleus*, bearing the positive charge, is responsible for almost the whole mass of the atom, while about it revolve negatively charged particles called *electrons*, whose total charge normally just neutralizes that of the nucleus. The mass of the electron, a common constituent in all atoms, has been shown by numerous and different methods to be about $\frac{1}{1846}$ of the mass of an hydrogen atom. The number of electrons, steadily increases one at a time, from 1 in hydrogen, 2 in helium, until we reach such numbers as 80 for mercury, and 92 for uranium. The hydrogen nucleus is sometimes called a *proton*, because there is much evidence that the nuclei of other atoms are built up of protons and electrons. For example, the nucleus of helium (atomic number 2) is considered to be built up of four protons and two electrons, thus giving it a net positive charge of two units. The number of positive charges on the nucleus, in terms of the electronic charge as unit, gives the *atomic number* of the element. This is the same as the number of electrons in the normal neutral atom.

By various means one or more electrons may be removed from the parent atom, thus leaving a positive *ion*. One such means, familiar to many in days of radio popularity, consists in,

(1) heating a metallic filament to incandescence in a highly evacuated vessel, (2) in having the filament negative with respect to a nearby electrode in the vessel. Suppose, for example, that *FF′*, Fig. 89, represents an incandescent filament, heated by being placed in an electric circuit with the battery *B*, and that it is surrounded by the metallic spiral *S*, the whole being in a highly exhausted vessel. If the filament is joined to the negative terminal of a second battery *A*, while the spiral is joined to the positive, electrons will pass from the filament to the spiral. Moreover,

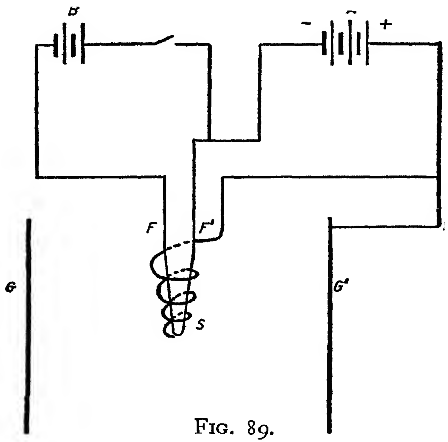


FIG. 89.

if an outer electrode *GG′* is electrically joined to the spiral so that no potential difference exists between *G* and *S*, then most of the electrons will pass through the spiral and traverse the outer region *with a speed controlled by the potential difference between the filament and the spiral.* Suppose, now, that the vapour of some metal is present in this region,—and this may readily be the case if the whole apparatus is heated sufficiently. The atoms of the metallic vapour will then be bombarded by the stream of electrons and so may be excited to luminosity. Obviously by varying the speed of the electrons, and so their mean kinetic energy, changes in the spectrum brought about by more and more violent

bombardment may readily be examined. Photographs of spectra taken in this way are reproduced in Fig. 1, Plate II. The single-line spectrum shown in *a* was obtained when the potential difference through which the electrons fell was 3.2 volts; the two-line spectrum in *b*, for a potential difference of 6.5 volts; while the many-line spectrum in *c* resulted from 10 volts. These photographs were taken by Drs. P. D. Foote, W. F. Meggers, and F. H. Mohler, through whose kindness it is possible to reproduce them in this book.

While the use of such a method requires much greater experimental skill than the others we have given, it has been thus briefly described because the results obtained by it are extremely important on account of the information provided concerning the origin of spectra.

(*g*) **The Electrodeless Discharge.** — Another source which the author has found extremely useful in the study of spectra is obtained by means of high frequency electrical currents. If a partially exhausted tube *without electrodes* is placed inside a coil carrying such currents, at suitable pressures a luminosity which is sometimes very brilliant takes place in the tube. By varying the pressure, or the intensity of the currents, changes in the spectra may be observed. Fig. 2, *b*, *c*, Plate II, and Fig. 4, Plate I, to which reference will again be made, give examples of spectra obtained in this way, with a tube containing mercury vapour.

High frequency currents may also be used to give bright luminous discharges in tubes to which *external* electrodes are applied. In this case the electrodes are joined to two points on a coil carrying the high frequency current.

69. Infra-Red and Ultra-Violet. — Before discussing information obtained by the use of the above light sources, a brief reference has to be made to another important fact.

Suppose the continuous spectrum **RV**, Fig. 90, of a white hot incandescent carbon rod is focussed on a screen by means described in the previous chapter. If the blackened bulb of a

sensitive thermometer is placed at places such as B in the region beyond the red end of the spectrum, a rise in temperature is observed. This is an indication of energy corresponding to invisible light, if we may use that term, in what is called the *infra-red* region. To explore this region carefully, delicate instruments capable of indicating the slight rise in temperature resulting from the absorption of an extremely small amount of energy are used. Of these we may mention, (1) the *thermopile* or the *thermocouple*, to which reference has already been made in Chapter I; (2) the

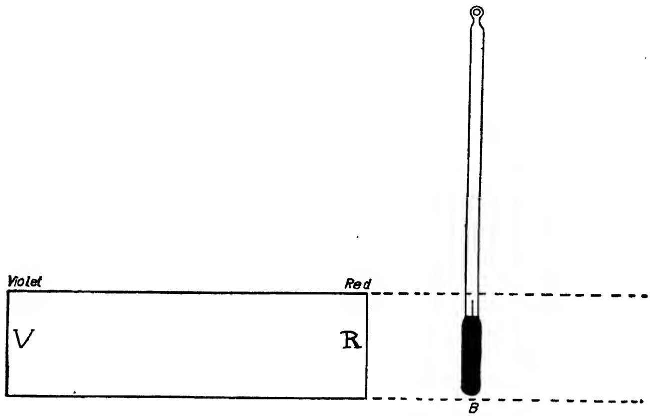


FIG. 90.

bolometer; (3) the *radiomicrometer*. The bolometer is an instrument in which a rise in temperature is indicated by a change in the electrical resistance of a blackened strip of platinum on which radiation is incident. As this strip forms one arm of a Wheatstone bridge, the increase in resistance with rising temperature causes a change in the deflection of the galvanometer in the circuit. This change in deflection is proportional to the intensity of the radiation falling on the strip. In the radiomicrometer, the radiant energy falls on a piece of blackened copper foil C, Fig. 91, which forms one juncture of an antimony-bismuth thermocouple. The thermo-electromotive force resulting from the rise

in temperature of the juncture causes a current in the circuit closed by a loop of wire suspended by a quartz fibre Q. Since this loop lies between the poles N and S of a magnet, a deflection results, just as in the case of a D'Arsonval galvanometer, whose magnitude is proportional to the current strength, and, therefore, to the rise in temperature, hence to the intensity of the incident radiant energy.

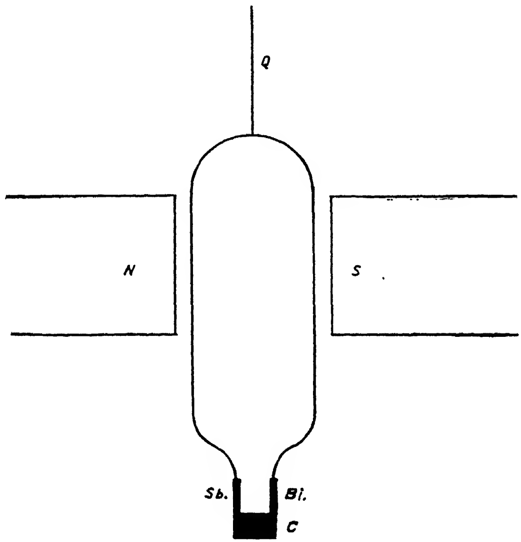


FIG. 91.

The general result of careful investigation in this region shows the existence of definite wave-lengths just as in the visible. For example, a mercury arc emits a strong wave-length λ10139 A, along with several others still longer. In special sources, wave-lengths in the extreme infra-red as long as 0.0412 cm. have been measured.

If a photographic plate suitably sensitized is substituted for the screen on which the spectrum is received, further evidence of the existence of the infra-red region can be obtained by the blacken-

ing of the plate in this region. In recent years great improvements have been made in the manufacture of plates sensitive to the infra-red and infra-red photography is now not confined to the examination of this part of the spectrum. Because infra-red rays penetrate hazy atmospheres much better than visible light, photographs showing marked detail may be taken through fog or at great distances. In infra-red photography green foliage and grass present the peculiar appearance of being like snow. This is due to the fact that chlorophyll reflects infra-red rays more strongly than visible light.

With any kind of a photographic plate, it is found that marked blackening occurs beyond the violet end of the visible spectrum. If, again, the ordinary spectrograph is replaced by one built on the same principle, but with lenses and prism made of quartz, the spectrum is found to extend far beyond the violet end. We have, therefore, still another region of invisible light, called the *ultra-violet*, a region in which, for most sources, isolated wave-lengths occur just as in the visible. For example, Fig. 4, Plate I, is a photograph of the spectrum of mercury vapour, taken with a quartz spectrograph, showing spectral lines extending to wave-lengths in the neighbourhood of 1900 angstroms. The spectra of Fig. 3, Plate I, and Fig. 2, Plate II, taken with a similar instrument, also show lines in the ultra-violet region.

With quartz prisms and lenses, wave-lengths as short as 1850 A can be measured, but beyond this limit quartz absorbs radiations to too great an extent to be of any use. The substitution of fluorite extends the limit as far as 1200 A while lithium fluoride is transparent to 1083 A. Wave-lengths as short as this bring us into the extreme ultra-violet region, where difficulties due to absorption become very great. The absorption of air itself, as well as that of the emulsion on the ordinary photographic plate, necessitates the use of special plates, and of instruments from which the air has been removed. To such an instrument reference will again be made, but at this stage we may note that wave-lengths as short as 40 A have been photographed in the extreme ultra-violet region.

The student should now realize, (1) that the visible spectrum is only a part of a much longer region of what may be called optical wave-lengths; (2) that in dealing with questions concerning the nature of light, it is necessary to consider facts pertaining to this more extended region. In other words we cannot confine our study of light to the restricted region of the visible spectrum. (See also Fig. 210.)

70. Emission Spectra. — We are now in a position to consider more carefully some of the results obtained by the use of such luminous sources as have been described. Since we are at present considering only the analysis of light *emitted* by a source, we are dealing with *emission spectra*. Of these we have two general classes, continuous and discontinuous. Continuous spectra, consisting of a wide range of unseparated wave-lengths, are emitted by incandescent solids, and, under very special circumstances, by fairly dense luminous gases.

Discontinuous spectra may be subdivided into two classes, (*a*) line spectra, (*b*) band spectra. Reference has already been made to the figures of Plate I, which provide good examples of line spectra, and attention need scarcely be directed to the sharply defined isolated wave-lengths which characterize them. Such spectra are emitted by substances in the gaseous or vapour state.

Sometimes the spectrum of a luminous source is characterized by groups of individual lines whose distance apart, never great, becomes less and less until frequently it is impossible to separate them. In Fig. 4, Plate II, the student will find a reproduction of such a group — a single band — taken by Dr. R. W. B. Pearse (England) with an instrument of high dispersion. With an instrument of low dispersion, frequently the individual members of a single band are not separated, so that, when one views a number of them, the spectrum has a fluted appearance, as illustrated by Fig. 3, Plate II, a portion of the band spectrum of nitrogen. The sharp edges where the individual bands are closest together are called *band heads*.

71. Plurality of Spectra. — A very little experience in using a spectroscope is sufficient to emphasize the fundamental fact of spectrum analysis, that every luminous source or substance emits a characteristic spectrum. Thus an observer with very little experience could tell at a glance that luminous hydrogen was the source in the case of Fig. 1, Plate I, luminous helium, in the case of Fig. 2, Plate I. As an extremely small amount of a substance in a luminous source is often sufficient to bring out its corresponding spectrum lines, the use of the spectroscope or spectrograph as a means of analysis is obvious.

Important as this fact is, it is still more important for the student to grasp the idea underlying the phrase *plurality of spectra*. By this is meant the fact that the same substance may give a variety of spectra when excited to luminosity by different means. Take the case of sodium, for example. If a little common salt or sodium itself is put in the non-luminous flame of a Bunsen burner, the resulting spectrum shows only a single yellow line, or, more accurately, with good dispersion and a narrow slit, two yellow lines close together. If, however, the sodium arc spectrum is photographed, many spectral lines are observed. Again, if the spark between two pieces of sodium is analyzed, some of the arc lines will be found to have disappeared, while certain new wave-lengths are present. We have then at least three different kinds of spectra, all characteristic of sodium, with some lines in common, but no two exactly alike.

In Fig. 2, Plate II, reproductions are given of three spectra of mercury vapour taken when the same tube was made luminous with different degrees of excitation. The increase in the number of lines as one passes from spectrum *a* to spectrum *c* is at once apparent. Reference has already been made to another example of plurality in the case of magnesium vapour when bombarded by electrons possessing varying amounts of kinetic energy. (See again Fig. 1, Plate II.)

In the early days of spectroscopy careful examination was made of the arc and the spark spectra of many elements. It was found that some lines, present with feeble intensity in the

arc, came out with greatly increased intensity in the spark. Such lines were said to be *enhanced*. In this connection note again Fig. 3, Plate I, where the spark spectrum of tin is superimposed on the arc spectrum, the two photographs being made on the same plate by allowing the light from the sources to fall on two different, but overlapping portions of the slit of the spectrograph. While allowance must be made for differences in exposure, the student should notice: (1) lines such as 2706, 2421, 2429 are present in both arc and spark, but weaker in the spark; (2) lines such as 2531, 2524, 2455, 2381 are present in arc, absent in spark; (3) lines such as 2658, 2643, 2631, absent in the arc, appear with marked intensity in the spark. The latter group are typical spark lines.

Because of modern theoretical developments, the term enhanced has fallen somewhat out of use. Anticipating later work, we may here state that spark lines are now known to be due to the atom in an ionized state, arc lines to the neutral atom. In the case of an element X, the arc spectrum is described as the X I spectrum, while X II refers to the spectrum of a singly ionized atom (one electron lost), X III to that of a doubly ionized atom (two electrons lost), and so on. In the case of chlorine, the atom of which may lose from one to seven electrons, lines corresponding to all the different ionized states have been observed.

72. Absorption Spectra. — If between the slit of a spectroscope and a source emitting a continuous range of wave-lengths a non-opaque substance is interposed, certain wave-lengths are sometimes absorbed, giving rise to corresponding dark regions in the spectrum. In such a case we are dealing with the *absorption spectrum* of the substance. As in the case of emission spectra, we may distinguish different kinds of absorption spectra.

A pure red glass, absorbing everything but red, gives a continuous absorption spectrum.

Dilute solutions of blood, of neodymium nitrate, and of many other substances give rise to isolated dark bands whose number, width, and position are characteristic of the absorbing substance.

This is a case of *selective* absorption, of which use may be made in the analysis of solutions of such substances.

In the case of an absorbing gas or vapour isolated wave-lengths, giving rise to lines as sharp as emission lines, may be observed. An excellent example of such an absorption spectrum is found when the sun is used as a source. With any instrument of even moderate dispersion and a narrow slit, it is easy to show that the spectrum of the sun is not continuous, but is crossed with narrow black lines. These lines are called

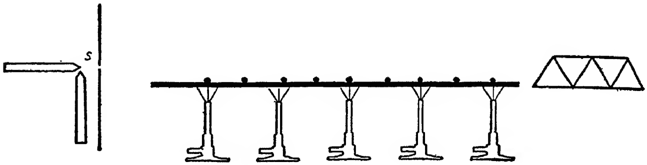


FIG. 92.

Fraunhofer lines after their discoverer, the German optician who made the first gratings and died in 1826 at the early age of thirty-nine.

In the early days of spectroscopy letters were assigned to a few of these lines. As an example, we may mention the A, C, D_1, D_2, F, and H wave-lengths of which some use was made in Chapter VI. The actual number of Fraunhofer lines is very great, as will be evident from an inspection of Figs 1 and 2, Plate III, reproductions of small portions only of the complete visible solar spectrum. The original photographs from which these figures were made were taken by Dr. R. E. DeLury of the Dominion Observatory, Ottawa, Canada, whose kindness has made possible their reproduction in this book.

By taking spectral photographs of the emission spectra of certain substances with the same instrument as was used to obtain solar spectra, a very important fact comes to light. *Some of the bright emission lines coincide with certain of the dark Fraunhofer lines.* For example, the two bright yellow lines of luminous

sodium correspond exactly to the dark lines called D_1 and D_2; two prominent lines in the spark spectrum of calcium correspond exactly to the dark lines called H and K.

The explanation will be found by a consideration of the following simple experiment. Suppose light, emerging from a slit illuminated by any intense white source S, Fig. 92, is allowed to pass through the column of vapour above pieces of fused salt or sodium strongly heated by burners below the plate on which the pieces rest. If the emergent light is examined by a spectroscope, the spectrum being thrown on a screen, it is found that in exactly the place normally occupied by the bright sodium line (or lines) a black image of the slit appears on the screen. This means that the sodium vapour absorbs just those wave-lengths which it emits when acting as a luminous source — a particular example of what is known as Kirchhoff's law. On the theoretical side we may, for the present, consider the phenomenon a case of resonance, the ultimate particles, whatever they are, absorbing the energy corresponding to each frequency in the continuous source which is the same as a natural frequency of vibration of the particles. (See, also, section 155, Chapter XVI.)

To return to the solar spectrum, we see that if a continuous spectrum is emitted by the hottest portion of the sun, and if enveloping this portion is a layer containing vapours of substances like sodium and calcium, we have a simple explanation of the presence of the D_1, D_2, H, and K (and other) Fraunhofer lines. A complete analysis of all the Fraunhofer lines reveals the existence of substances present in the solar atmosphere.

When, due to absorption, the absence of a bright spectral line is marked by a dark line, as in the sodium experiment, the line is said to be *reversed*. All Fraunhofer lines, therefore, are examples of *reversals* of corresponding emission lines.

A special case of reversal is found in what is known as the *self-reversal* of a line. This happens when a spectral line appears bright, although frequently broader than usual, with a narrow dark centre. If the vapour in an arc, for example, is sufficiently dense, the spectral line may be broadened, while the cooler, less

dense vapour in the outer portions of the arc absorbs the wave-length corresponding to the centre of the line.

73. Spectral Series. — Most students are familiar with the fact that when an organ pipe is blown, the note frequently consists of the fundamental, together with the harmonic overtones, whose frequencies are simple integral multiples of the fundamental. It was not unnatural for men to look for some similar simple relationship obtaining among the various wave-lengths emitted by a luminous source. Indeed, the regular spacing of the important spectral lines in the spectrum of hydrogen (Fig. 1, Plate II)

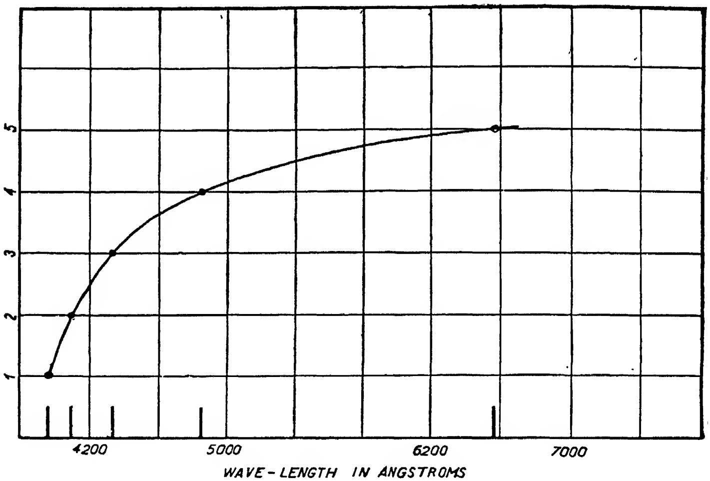


FIG. 93.

strongly suggests the possible existence of some such relationship. That there is some law connecting these lines becomes practically certain when one makes a plot similar to that of Fig. 93 (where one ordinate represents wave-lengths, the other any sequence of integral numbers), and finds that a smooth curve is the result.

In 1885 it was shown by Balmer (Germany) that the wave-lengths of these hydrogen lines could all be calculated from the relation

$$\lambda = 3645.6 \frac{m^2}{m^2 -} \tag{7.01}$$

where m takes values 3, 4, 5 . . ., and λ is in angstroms.
For reasons which will appear later, this relation may more conveniently and accurately be expressed

$$\bar{\nu} = \frac{1}{\lambda} = 109678 \left(\frac{1}{2^2} - \frac{1}{m^2} \right), \tag{7.02}$$

where $\bar{\nu}$, the number of wave-lengths in 1 cm., is called the *wave-number*. In (7.02) λ is in centimetres.

Since $c = \nu\lambda$, where c is the velocity of light in free space and ν is the frequency of the light, we have $\nu = c\bar{\nu}$.

In dealing with wave-numbers, note that wave-lengths *in vacuo* are always used. To change from λ *in vacuo* to λ in I.A. (see p. 133) the index of refraction of air at 15° C. and 760 mm. pressure must be known, since $\lambda_{vac.} = \text{index} \times \lambda_{air}$. Books giving the wave-lengths of spectral series lines, however, usually include a table of conversion values.

In Table VIII a comparison is made of a few hydrogen wave-lengths (expressed in I.A.) as calculated by relation (7.02) with values observed.

Table VIII

A Few Wave-lengths of the Balmer Hydrogen Series

	Observed	Calculated
H_α	6562.84	6562.79
H_β	4861.36	4861.33
H_γ	4340.48	4340.47
H_δ	4101.76	4101.74

A group of wave-lengths, any one of which may be calculated by using the appropriate *integral* value of the variable m, constitutes a *spectral series*. The above hydrogen lines constitute the first four members of what is usually called the Balmer hydrogen

series. As many as thirty-five members of this series have been observed, that is, as far as $m = 37$.

It may not be amiss at this place to point out that the Balmer formula, like all other series relations, involves, not the first power of the variable m, but the second power.

Since the pioneer work of Balmer, a tremendous amount of work has been done in the analysis of spectra, with the result that not only have spectral series been discovered for most elements but, also, for one and the same element, several different series are known.

In hydrogen, for example, there is another group of lines in the ultra-violet whose wave-lengths may be calculated from the relation,

$$\bar{\nu} = 109678\left(\frac{1}{1^2} - \frac{1}{m^2}\right), \qquad (7.03)$$

where m takes values 2, 3, 4 . . . . This group constitutes the Lyman series named after its discoverer, Prof. Lyman, an American physicist. The student can easily show by calculation that lines observed at 1216, 1026 and 972 constitute the first three members of this series.

In the infra-red region three other hydrogen series have been observed. As these are of the same kind as the Balmer and the Lyman series, they may be represented by the general formula

$$\nu = 109678\left(\frac{1}{k^2} - \frac{1}{m^2}\right). \qquad (7.04)$$

When $k = 3$ and $m = 4, 5, 6, \cdot \cdot \cdot$, we have the Paschen series. The observed wave-lengths range from 18751($m = 4$) to 8863($m = 11$).

When $k = 4$ and $m = 5, 6, 7, \cdot \cdot \cdot$, the Brackett series is obtained. This includes lines with wave-lengths as long as 40500.

Finally, when $k = 5$ and $m = 6, 7, \cdot \cdot \cdot$, there is the Pfund series, in which only a single member has been observed. Its wave-length is 74600.

The number 109678 which occurs in all relations, and will appear in all other series relations, is called the Rydberg constant, and will be represented by N.*

Graphical Representation of Hydrogen Series. — Suppose, in Fig. 94, that the distances between the horizontal lines or levels from an arbitrary zero line AB are proportional to the magnitudes $\frac{N}{1^2}, \frac{N}{2^2}, \frac{N}{3^2}, \frac{N}{4^2} \cdot \cdot \cdot \cdot$ Then the Balmer series may be represented by the vertical lines joining the $\frac{N}{3^2}, \frac{N}{4^2}, \frac{N}{5^2} \cdot$ levels to the $\frac{N}{2^2}$ level; the Lyman series by the group of vertical lines joining $\frac{N}{2^2}$, $\frac{N}{3^2} \cdot \cdot \cdot$ levels to $\frac{N}{1^2}$; and the Paschen series by the lines joining $\frac{N}{4^2}$, N to $\frac{N}{3^2}$. The advantage of thus representing spectral series will appear when questions relating to the origin of spectra are under consideration.

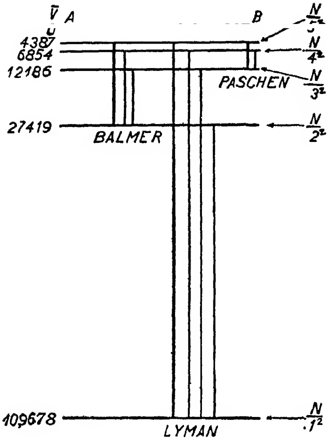


FIG. 94.

General Series Relations. — Some of the more important ideas in connection with series for elements other than hydrogen may be explained by taking the case of lithium. In this element, there are three important series of arc lines.

(a) The *principal* series,

$$\nu_P = 43486 - \frac{N}{(m + 0.9597)^2}, \quad m = 1, 2, 3, \qquad (7.05a)$$

* Strictly speaking this number varies slightly in different series relations for reasons which will be apparent in Chap. XVIII.

(*b*) The *diffuse* series,

$$\bar{\nu}_D = 28582 - \frac{N}{(m + 0.9990)^2}, \ m = 2, 3, 4, \tag{7.05b}$$

(*c*) The *sharp* series,

$$\bar{\nu}_S = 28582 - \frac{N}{(m + 0.5888)^2}, \ m = 1, 2, 3, \tag{7.05c}$$

As in hydrogen, the wave-number for each wave-length is equal to the difference between a fixed number, called the *limit*, and a variable *term*, which involves the universal Rydberg constant N and the variable *sequence* number m. It is well to note several other things.

(*a*) The variable term is no longer of the simple form $\frac{N}{m^2}$, but $\frac{N}{(m + c)^2}$, where c is a constant whose value varies with the particular series under consideration.

(*b*) The same number gives the limit of the sharp and of the diffuse series. This is indicated in Fig. 95, where the dotted lines

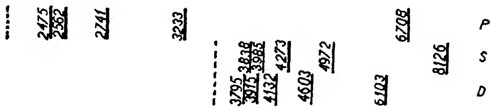


FIG. 95.

represent limits and where an individual wave-length is indicated by drawing a short vertical line on a horizontal *wave-number* scale.

(*c*) The limit of the principal series, namely, 43486 is numerically equal to the first term of the sharp series. Thus when $m = 1$,

$$\frac{109678}{(m + 0.5888)^2} = 43486.$$

The common limit of the sharp and the diffuse series is equal to the first term of the principal. Again, when $m = 1$,

$$\frac{109678}{(m + 0.9597)} = 28582.$$

These are particular examples of a very important general law, which states that the limit of any series is a term of some other series.

Bearing the above result in mind, we may now re-write the above expressions,

$$\bar{\nu}_P = \frac{N}{(1 + 0.5888)^2} - \frac{N}{(m + 0.9597)^2} \tag{7.06a}$$

$$\bar{\nu}_D = \frac{N}{(1 + 0.9597)^2} - \frac{N}{(m + 0.9990)^2} \tag{7.06b}$$

$$\bar{\nu}_S = \frac{N}{(1 + 0.9597)^2} - \frac{N}{(m + 0.5888)^2}. \tag{7.06c}$$

We see, then, that the wave-number of any spectral line (which is a member of a series) is equal to the difference between two terms. This is a statement of the Ritz *combination* law. For convenience in writing, the function $\frac{N}{(m + c)^2}$ is written mc, in consequence of which we may write the series relations as follows:

$$\bar{\nu}_P = 1S - mP \tag{7.07a}$$

$$\bar{\nu}_D = 1P - mD \tag{7.07b}$$

$$\bar{\nu}_S = 1P - mS, \tag{7.07c}$$

where P, D and S are used to represent the constants of the principal, diffuse and sharp series.

By means of the same method as used for hydrogen, these three series have been represented graphically in Fig. 96. The second column represents the P term values or levels, that is

$$\frac{N}{(1 + 0.9597)^2}, \quad \frac{N}{(2 + 0.9597)^2}, \quad \frac{N}{(3 + 0.9597)^2}, \ \cdots; \text{ the first}$$

column, the S terms; and the third, the D terms, while as before lines connecting levels represent wave-lengths in accordance with relations (7.07). The student should not have difficulty in interpreting the diagram for himself. (See also Chapter XIX.)

The terms principal, sharp, and diffuse were originally chosen because of certain physical differences in the lines which belonged to the individual series. Thus, members of principal series were strong lines, members of the diffuse were lines with ill-defined edges, while sharp series lines were sharply defined. As this is

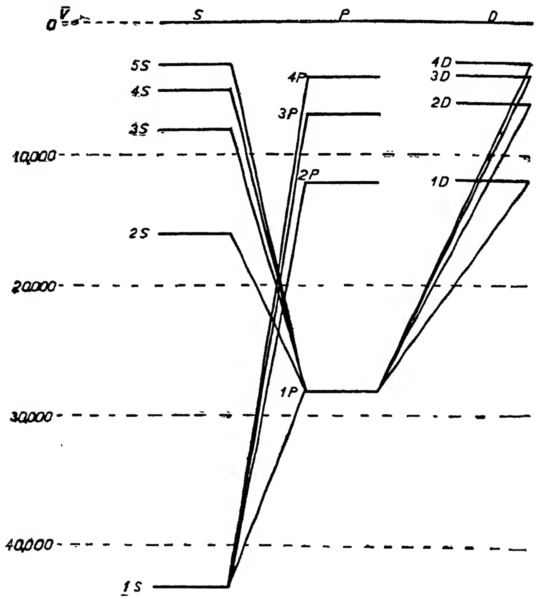


FIG. 96.

not always the case, however, and as still other types of series are known, the only exact description of any series is to be found in the relation which describes it, as well as in its connection with other series.

Doublet, Triplet and Multiplet Series. — The study of series is intimately bound up with another important fact which must now be explained. The student's attention has already been directed to the two yellow sodium lines (D_1 and D_2) in the flame

spectrum. Now a glance at the arc spectrum of this element shows the presence of other pairs of lines. For many of these pairs the difference in the wave-numbers of the two lines is exactly the same, while for other pairs this difference becomes less, the farther one goes in the ultra-violet. Such pairs of lines are called *doublets*. Similar pairs occur in all the members of the sodium family, although in lithium the lines are so close that they appear unresolved in ordinary spectrographs. Now these doublet lines are all members of series. Thus in the sodium arc spectrum we

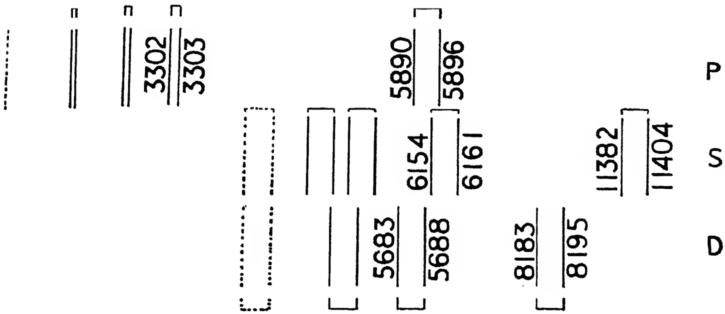


FIG. 96A.

have six important series: — two principal, the first members of which are the D lines, two diffuse and two sharp. The doublet separation is the same for all pairs belonging to the diffuse and the sharp series, thus giving rise to parallel series, while, in the case of the principal series, the doublet separation gradually decreases until the two series reach a common limit.

Because of the comparatively small separation of the doublet pairs in the sodium series, it is not possible to represent them to scale by a diagram similar to Fig. 95. In Fig. 96A, however, the general nature of the important sodium series has been indicated by assuming a much larger separation between the doublets than that which actually exists.

In other elements there are related lines appearing in triplets. In still others we may have quadruplets, quintets, etc.

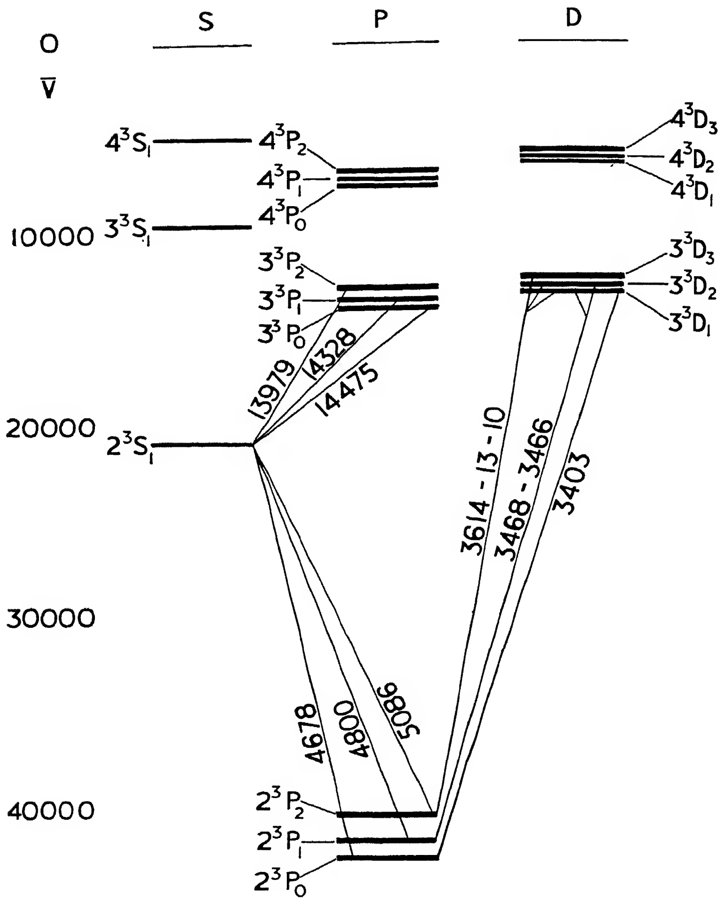


FIG. 96B.

73A. On page 169 reference has been made to the Ritz combination law according to which the wave-number of a spectral line is equal to the difference between two terms. In one sense

term values, therefore, are more fundamental than the wave-numbers of *lines*, and the analysis of a spectrum into series consists essentially in establishing term values, that is, in constructing diagrams like Fig. 96. The question now arises, what alteration must be made in such diagrams when we pass from singlet to multiplet series? Although a knowledge of the theoretical significance of spectral series is necessary before this question can be satisfactorily answered, at this stage we may indicate the chief change that is necessary, postponing further discussion until section 168A.

As a concrete case, consider the principal, sharp and diffuse triplet series lines of cadmium, some of which are marked in Fig. 1, Plate VII. The wave-numbers of all these lines can be represented as in Fig. 96B (drawn roughly to scale) by a series of single S term values, which we may designate mS, with m having successive integral values, just as in Fig. 96; a series of P term values, each of which consists of three sub-terms; and a series of D terms, each of which has also three sub-terms. To distinguish sub-terms the notation of page 169 must be extended. This is done by the use of subscripts and superscripts. Thus, in Fig. 96B, the lowest group of P terms is designated 2^3P_0, 2^3P_1, and 2^3P_2. Here the serial number m has the value 2 and the three sub-levels are distinguished by the subscripts 0, 1, 2 at the lower right of P. The superscript 3 to the upper left of P is used simply to indicate that we are dealing with a triplet system. The next P group is designated 3^3P_0, 3^3P_1, 3^3P_2; the lowest D group 3^3D_1, 3^3D_2, 3^3D_3 and so on. Although the S terms are single, since they are part of the whole triplet system they are usually marked m^3S_1. The first

A Sharp Triplet of Cadmium

Wave-length	$\bar{\nu}$	Notation
5085.88	19656.8	$2^3P_2 - 2^3S_1$
4799.91	20827.9	$2^3P_1 - 2^3S_1$
4678.19	21369.8	$2^3P_0 - 2^3S_1$

three members of the sharp series of cadmium, that is, the lines whose wave-lengths are 5086, 4800, 4678 (see Fig. 1, Plate VII) are therefore described as on previous page.

Since both the P and the D terms occur in groups of three sub-terms, we might expect diffuse lines to occur in groups of nine. This would certainly be the case if each of the three D sub-terms combined with each of the three P sub-terms. Actually the D lines occur in groups of six, because a selection principle, which we shall discuss later (section 168A), rules out three of the possible combinations. The first group of six D lines and their designation are given below.

A Diffuse Triplet of Cadmium

Wave-length	$\bar{\nu}$	Notation
3614.43	27659.0	$2^3P_2 - 3^3D_1$
3612.89	27670.8	$2^3P_2 - 3^3D_2$
3610.51	27689.0	$2^3P_2 - 3^3D_3$
3467.61	28830.1	$2^3P_1 - 3^3D_1$
3466.18	28842.0	$2^3P_1 - 3^3D_2$
3403.60	29372.3	$2^3P_0 - 3^3D_1$

In Fig. 1, Plate VII these six lines appear as three, because the dispersion of the spectrograph used in making the spectrogram was not great enough to separate 3614, 3613 and 3610 nor 3467.6 and 3466.2.

In cadmium, in addition to the above triplet system of lines, there are many lines which belong to a singlet system. That is, we have a singlet S series, a singlet P series, a singlet D series, etc. and corresponding to these lines, a series of single S terms (marked m^1S_0), a series of single P term (marked m^1P_1) and a series of single D terms (m^1D_2). To avoid confusion, the singlet system term values have not been included in Fig. 96B.

The same ideas apply to the term analysis of a doublet system like that of sodium and the other alkali elements, where the wave-numbers of the S, P and D series lines can all be found from a series of single S term values; a series of double-valued P terms,

and a series of double-valued D terms.* To distinguish the sub-levels of the P terms the subscripts $\frac{3}{2}$ and $\frac{1}{2}$ are used, whereas $\frac{5}{2}$ and $\frac{3}{2}$ are used to distinguish the D sub-levels. The D lines of sodium, that is, 5896 and 5890, the first members of the doublet P series, are described by

$$1\,{}^2S_{\frac{1}{2}} - 1\,{}^2P_{\frac{1}{2}} \quad \text{and} \quad 1\,{}^2S_{\frac{1}{2}} - 1\,{}^2P_{\frac{3}{2}}.\dagger$$

To represent completely, therefore, all the S, P and D doublet series of sodium, Fig. 96 should be extended in the same manner as has been done for a triplet system in Fig. 96B. In the doublet system all S terms would remain single as in Fig. 96 but, in accordance with the notation already explained, would now be marked $m^2S_{\frac{1}{2}}$ (where $m = 1, 2, 3$ etc.); the P terms would all occur in pairs marked $m^2P_{\frac{3}{2}}$ and $m^2P_{\frac{1}{2}}$; and the D terms in pairs marked $m^2D_{\frac{5}{2}}$ and $m^2D_{\frac{3}{2}}$. In the case of the element lithium, the doublet lines are all so close that the sub-terms appear as one in such diagrams; and in sodium, the D sub-terms are so close together that they also appear as one. In passing it is interesting to note that in a chemical family like the alkalies, all of which give doublet spectra, the doublet separation steadily increases with increasing atomic weight. For example, if we take the separation (in wave-numbers) of the first pair of the principal series, we have 0.34 for lithium, 17.2 for sodium, 57.8 for potassium, 237 for rubidium, and 554 for caesium.

For convenience, the designation of sharp, principal and diffuse terms, for singlet, doublet and triplet systems has been summarized in Table B.

With respect to a singlet system it will be noted that this notation (which is now used internationally) differs slightly from that used in section 74. In accordance with this more modern usage, relations 7.07*a*, *b* and *c* would be written

$$\begin{aligned} \bar{\nu}_P &= 1\,{}^1S_0 - m^1P_1 \\ \bar{\nu}_D &= 1\,{}^1P_1 - m^1D_2 \\ \bar{\nu}_S &= 1\,{}^1P_1 - m^1S_0. \end{aligned}$$

* In some elements these double term values are too close to be separated.

† The superscript 2 at the upper left of the letter indicates the doublet nature of the system.

Table B

Designation of Terms

	Sharp	Principal	Diffuse
singlet	m^1S_0	m^1P_1	m^1D_2
doublet	$m^2S_{\frac{1}{2}}$	$m^2P_{\frac{3}{2}}$ $m^2P_{\frac{1}{2}}$	$m^2D_{\frac{5}{2}}$ $m^2D_{\frac{3}{2}}$
triplet	m^3S_1	m^3P_2 m^3P_1 m^3P_0	m^3D_3 m^3D_2 m^3D_1

73B. The general analysis which has been made of the spectra of elements shows (1) in any system of series for a single element, we have often to consider more *types* than the sharp, principal, and diffuse. These are designated F, G, H, etc. (2) In addition to singlet, doublet and triplet systems, we have also, for the more complex spectra, quadruplet, quintet, sextet, etc. systems. Further discussion of such details is beyond the scope of this book and we shall conclude this section with a statement of two other interesting facts. In a given chemical family, for all elements the arc spectra are of either even or odd multiplicity — never both, but as we pass from one vertical column in the periodic table to the next, the character of the group alternates. Thus in the sodium family, doublets only occur in arc spectra; in the magnesium, triplets and singlets.

74. In the remainder of this chapter attention will be directed to certain means by which a single wave-length has its appearance or position altered, or is replaced by more than one line.

Doppler Effect. — It is taken for granted that readers of this book are familiar with the fact that when a source, emitting a note of a certain pitch, is approaching an observer, he hears a note of higher pitch, whereas, if the source is receding, the note heard is of lower pitch. A similar phenomenon is observed in light.

Suppose a vibrating source, from which waves of any kind are spreading out, is moving with velocity v in the direction of

the arrow, Fig. 97. Suppose, further, that at a certain instant, the source is at the position marked 0, that T secs. before it was at 1, $2T$ secs. before at 2, $3T$ secs. before at 3, and so on. In the diagram an attempt has been made to represent, at the particular moment that the source is at 0, the positions of wave-fronts which left the source when it was in positions 1, 2, 3, etc. The wave-front which left the source when at 1 will have advanced to the position $AA'A''$, where $1A = VT$, V being the velocity

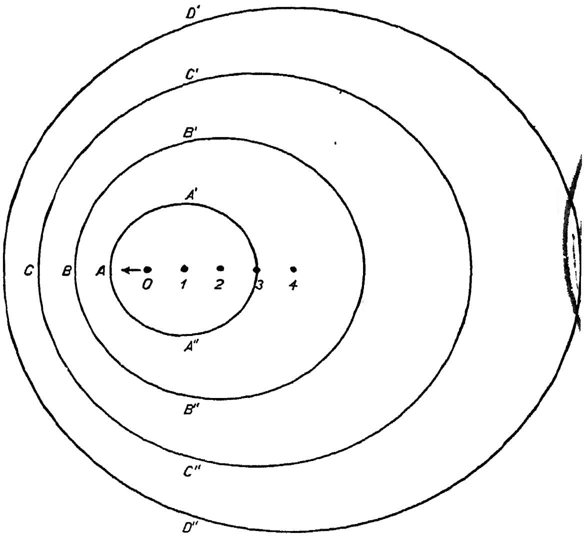


FIG. 97.

of the wave disturbance; the wave-front originating at 2 will have reached $BB'B''$, where $2B = 2VT$. Similarly, $3C = 3VT$, $4D = 4VT$. If T is the periodic time, then these wave-fronts are a wave-length apart.

Obviously, then, the wave-fronts are crowded more closely together in the direction towards which the source is moving, while they are farther apart than normally in the opposite direction. If the source is a luminous one and we accept light as a

wave phenomenon, then an observer towards which the source is moving will receive wave-lengths shorter, an observer from which the source is receding, wave-lengths longer, than those emitted by the source if stationary. The actual change in wave-length is readily found.

Let λ' = wave-length from an approaching source,

λ_0 = wave-length from a stationary source.

Then, since $oD = 4\lambda'$

and $oD = 4D - o4$

$$\therefore \quad 4\lambda' = 4D - o4$$
$$= 4VT - 4vT$$

or

$$\lambda' = VT\left(1 - \frac{v}{V}\right)$$

$$= \lambda_0\left(1 - \frac{v}{V}\right). \qquad (7.08a)$$

For a receding source,

$$\lambda' = \lambda_0\left(1 + \frac{v}{V}\right). \qquad (7.08b)$$

Such changes in wave-length are observed with moving light sources, but, because of the high value of V, the actual displacement of any spectral line is extremely small. To realize this, take a concrete case. To obtain a displacement of one-tenth of an angstrom unit for light of wave-length 5000 angstroms, the application of relation (7.08) shows that a source would have to move with a velocity of 5 kilometres per sec. (remember that $V = 3 \times 10^{10}$ cm. per sec.), or about 1,340 miles per hour. Such a displacement is about one-sixtieth of the separation of the D lines of sodium. But small as it is, much smaller displacements can be measured, and in many astronomical observatories, measurements of Doppler displacements, and subsequent calculation of velocities are part of the regular routine work. As examples of phenomena on which such observations are made, we note briefly the following:

(*a*) **The Mapping of the Path of Certain Stars in the Heavens.** — By photographing the spectrum of the light from a star, night after night, any possible motion of the star toward or away from the earth is made evident by the corresponding shift in the position of any line occurring in the spectrum. It is found, for example, that for a period of a few months a spectral line may show a displacement toward the violet end of the spectrum, the amount of which gradually decreases until finally it drops to zero. Continued observation then shows a gradually increasing displacement toward the red end. Such a star is at first approaching the earth, afterwards receding from it.

(*b*) **Measurements Relating to the Rotation of the Sun.** — The two spectra of Fig. 2, Plate III, provide an excellent example of such observations. To obtain spectrum *a* (only a small portion of the complete visible region), light from the west limb of the sun, at the equator, fell on a portion of the slit of the instrument, while, for spectrum *b*, light from the east limb fell on an adjacent portion of the slit. Because of the rotation of the sun, one limb is approaching the earth, the other receding. Hence in one case there is a slight shortening of each wave-length, for the other a slight lengthening. The resulting displacement of corresponding lines in the two spectra, while not of great magnitude, is clearly shown in the figure. Actual measurement (obtained by placing the photographic plate under a high-power travelling microscope) shows that this displacement corresponds to a value of somewhat less than 2 kilometres per second.

(*c*) **Investigation of Sun-Spots.** — Fig. 1, Plate III, is a photograph of a portion of the spectrum obtained when the image of a sun-spot is formed on one part of the slit of the instrument, while the adjacent parts are illuminated by light from the normal surface of the sun outside the spot. The dark band across the spectrum corresponds to the centre of the spot, the *umbra*. In passing from the umbra to the normal surface we go through the *penumbra*. A close examination of the photograph will show a waviness in any individual line as one goes through the sun-

spot region. This waviness is the result of a Doppler displacement of the lines in the penumbra, in opposite directions on each side of the centre, somewhat as represented in *a*, Fig. 98. This displacement, to quote Dr. DeLury, to whom the author is greatly indebted both for the photographs and for information relating to them, is "apparently caused by motions of the gases approximately tangential to the solar surface outward from the centre of the spot for most of the solar lines, but zero for some lines at higher levels, and inward for some lines at still higher levels." Thus most of the lines in the photograph have the displacements

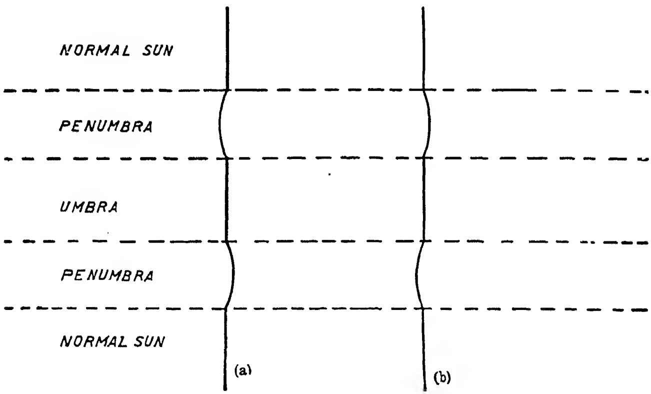


FIG. 98.

corresponding to motions of the first type, or are represented by *a*, Fig. 98, but a close examination will show that the line marked *H* (calcium, λ 3968 A) is an example of the less common higher level type, for which the displacement is represented by *b*, Fig. 98.

75. The Zeeman Effect. — When a source of light is placed between the poles of a powerful electromagnet, it is found that, when the magnetic field is thrown on, a single spectral line is replaced by a number of others, that is, breaks up into components. This separation resulting from the action of a magnetic field on the source is called the *Zeeman Effect*, a name given in

PLATE III

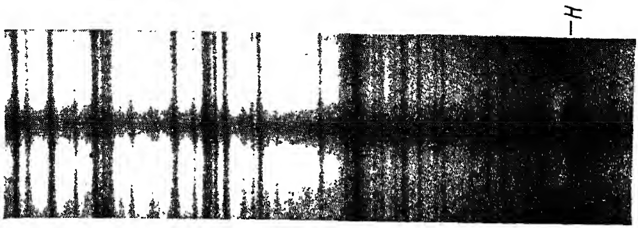


FIG. 1. A portion of the spectrum of the sun in the neighborhood of a sunspot. (DeLury)

FIG. 2. A portion of spectrum of east limb of the sun side by side with corresponding portion of west limb. To show displacement of lines due to Doppler effect. (DeLury)

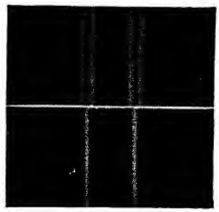


FIG. 3. Normal Zeeman Effect. (McLennan)

FIG. 4. Zeeman effect. (*a*) magnetic field off; (*b*) field on, vision in direction of field. (McLennan)

FIG. 5. Anomalous Dispersion of sodium vapour.

honour of its discoverer, Prof. P. Zeeman, of Holland. The number of components depends on the spectral line examined, and when the spectroscope is directed in a line at right angles to the lines of force, the result is not the same as when the direction is along the lines of force. Further discussion of this phenomenon will be given in Chapter XVIII, but at this point attention may be directed to Fig. 3 and Fig. 4, Plate III. Figure 3, an example of the so-called normal Zeeman effect (for zinc $\lambda 6362$), shows the characteristic group of three when observation is at right angles to the lines of force (the transverse position). Figure 4 is an example of longitudinal observation, and shows how the single line in Fig. *a*, where there is no magnetic field, is replaced in *b*, where the field is applied, by two components, one of either side of the normal position of the line. It may be of interest to note that the line in this case is a somewhat famous one, being a green one of wave-length $\lambda 5577$, which is observed in the auroral spectrum. An extended study of this line has been made by Dr. J. C. McLennan, F.R.S., of the University of Toronto, Canada, through whose kindness it has been possible to reproduce these photographs.

76. Stark Effect. — As might be expected, an electric as well as a magnetic field has an influence on the light emitted by a luminous source. If the electric field is sufficiently powerful, a single line is replaced by components as was first shown by J. Stark (Germany). A good example of the Stark effect is shown in Fig. 2, Plate VI, where the single line at the top of the spectrum is replaced (at the bottom, in the region of the strong electric field) by a number of components. For the original photograph the author is indebted to Dr. J. Stuart Foster, of McGill University, Montreal, a physicist who has made an extended study of this phenomenon.

While a discussion of the theoretical significance of the phenomena described in this chapter can best be given at a later stage, the mere enumeration of the facts themselves should be sufficient to make the reader realize what a marvellous instrument

the spectroscope is. By means of it we can identify substances in the laboratory and learn much of their ultimate structure; we acquire a vast amount of knowledge regarding the composition and motion of heavenly bodies; and we learn of the existence of law and order in atomic microcosms, as well as in the macrocosm.

It is remarkable that the great Newton, who gave the world the universal law of gravitational attraction, valid in the laboratory as well as throughout the solar system, should also be the pioneer with whose name we associate the prism and the spectrum.

Problems

1. The displacement of the F line of hydrogen (wave-length 4861 angstrom units) in the spectrum of a star is 0.1 of a unit toward the violet. What are the direction of motion and the velocity of the star?

2. How many vibrations per second will be received from a bicycle whistle giving out 500 vibrations per sec. and approaching at the rate of 10 miles per hour?

3. Describe the flame spectrum of sodium.

4. *The limit of a spectral series is a term of some other series.* Discuss this statement with reference to the important lithium series.

5. A spectral line is represented by the formula

$$\bar{\nu} = 2P - 3S,$$

where $P = 0.62$, $S = 1.51$.

(*a*) Find its wave-length. (*b*) Find the wave-number giving the limit of the series to which the line belongs. (*c*) Find the value of the term $1S$.

6. Use the Balmer series relation for hydrogen,

$$\bar{\nu} = N\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$$

to calculate the wave-lengths of the first five members of the series.

CHAPTER VIII

INTERFERENCE

In Chapter I it was pointed out that the wave-theory of light received general acceptance only after Thos. Young had demonstrated the phenomenon of interference. In the two chapters preceding this one, frequent use has been made of the numerical magnitudes of wave-lengths, but as yet no method has been described for actually measuring wave-lengths. The purpose of this chapter is to show how a detailed consideration of interference leads us to various means of evaluating wave-lengths, as well as to many other important applications.

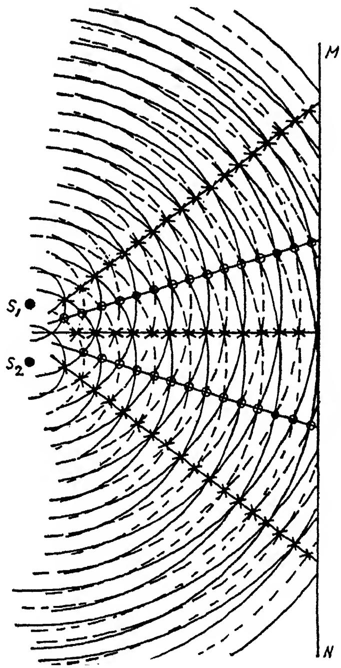


FIG. 99.

77. Suppose S_1 and S_2, Fig. 99, are two sources vibrating with equal amplitudes, with the same periods, and sending out waves in any continuous medium. The figure has been drawn to represent, at a certain instant, the position of regions of maximum and of minimum displacement, that is, of what we may, by analogy with water waves, call crests and troughs. The maxima are represented by the unbroken lines, minima by dotted lines. An examination of the figure will show that at certain places, marked ×, crests and crests, or troughs and troughs occur simultaneously;

at other places, marked o, crests and troughs are superimposed. Now, if we apply the principle of superposition, it follows that at the former places the resultant displacement is double that due to either disturbance separately, whereas at the latter it is zero. Lines joining points marked × are therefore loci of particles for which there is a marked resultant disturbance, while lines joining the other set of points are loci of particles which are not displaced at all.

If S_1 and S_2 represent disturbances caused by the vibrations of the two prongs of a low frequency tuning fork disturbing the

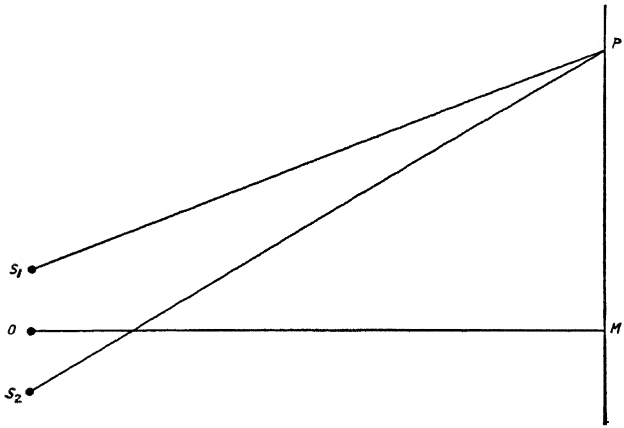


FIG. 100.

surface of water or of mercury, the existence of such lines can readily be shown. If S_1 and S_2 are luminous sources, and a screen is placed at MN, it follows that on the screen we should have places of maximum displacement alternating with others of zero displacement, — that is, alternately light and darkness.

For a quantitative discussion of this phenomenon we shall approach the question somewhat differently.

Consider any particle P, Fig. 100, which is disturbed as a result of wave-motion coming from the two sources S_1 and S_2, vibrating with equal amplitudes, with the same period, and with

no phase difference between them. Our problem is to find the resultant disturbance at P. This we shall do by two different methods, one graphical or semi-graphical, the other analytical.

Method I. — In this method we can make use of the proposition proved in section 15, Chapter II, for, whatever the position of P, we have to do with the resultant of two *S.H.M.* of the same period. The resultant amplitude is therefore represented by $MN = A$, Fig. 101, where $MO = a = ON$ represents the amplitude of P due to each disturbance separately, and δ is the difference in phase of the two disturbances on reaching P. Obviously δ will depend on the position of P, in Fig. 100, and will continuously increase the greater the distance of P from the point M, where

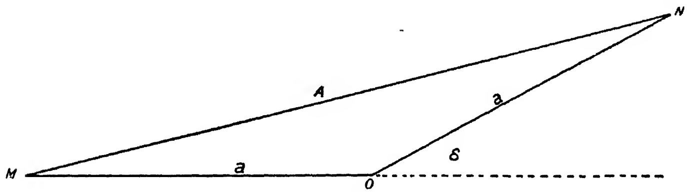


FIG. 101.

$S_1M = S_2M$, and PM is perpendicular to OM. The resultant amplitude, therefore, will also continuously change with the position of P, going through a cycle of values. Thus, when $\delta = 0$, A is a maximum $= 2a$; as δ increases, A decreases until when $\delta = \pi$, $A = 0$; for still greater values of δ, A increases, reaching a second maximum equal to $2a$, when $\delta = 2\pi$. The cycle is then repeated and we see that, if we had two such light sources as S_1 and S_2, the resultant disturbance at points on the screen should regularly alternate from maximum to minimum values, that is, from light to darkness.

By simple geometry,

$$\begin{aligned} A^2 &= a^2 + a^2 + 2a^2 \cos \delta\,^* \\ &= 2a^2(1 + \cos \delta), \end{aligned} \tag{8.01}$$

* If the two disturbances are of unequal amplitudes, a and b,

$$A^2 = a^2 + b^2 + 2ab \cos \delta.$$

which again tells us that whenever δ is increased by 2π, values of A^2 repeat, being equal to $4a^2$ when $\delta = 0, 2\pi, 4\pi, 6\pi \cdot \cdot \cdot$, that is, when $\delta = 0$ or $n \cdot 2\pi$, where n is integral.

Now, as already proved in section 19, Chapter II,

$$\text{phase difference} = \frac{2\pi}{\lambda} \cdot \text{path difference},$$

$$\therefore \qquad = \frac{2\pi}{\lambda} \cdot (S_2P - S_1P), \text{ for Fig. 100.}$$

$\therefore$ at maximum values of A^2,

$$n \cdot 2\pi = \frac{2\pi}{\lambda} \cdot (S_2P - S_1P)$$

or
$$S_2P - S_1P = n \cdot \lambda$$
or
$$S_2P - S_1P = 0, \text{ when } \delta = 0.$$

In other words, whenever the path difference is equal to a whole number of wave-lengths (or is zero) maximum values of the resultant disturbance exist. In the same way it can readily be shown that for minimum values, the path difference is equal to an odd number of half wave-lengths.

If the phase difference between S_1 and S_2 is not zero (although always a constant amount), the only difference in the above solution is that the central maximum does not occur at M. Just as before, however, whenever the path length is increased or decreased by one wave-length, we go from one maximum or one minimum to the next.

Since the intensity is proportional to the square of the amplitude, a plot of $2a^2(1 + \cos \delta)$ for various values of δ will show how the intensity varies as we go through maxima and minima values. Such a plot is given in Fig. 102. It should be noted that the width of bright regions (maxima) is about the same as that of dark places (minima).

Method II.

Let

$$y = a \sin\left(\frac{2\pi t}{T} - \alpha_1\right)$$

represent the disturbance reaching P from S_1,

and
$$= a \sin\left(\frac{2\pi t}{T} - \alpha_2\right)$$

represent that reaching P from S_2.

It follows that $\alpha_1 - \alpha_2 = \delta$, the phase difference between the two disturbances.

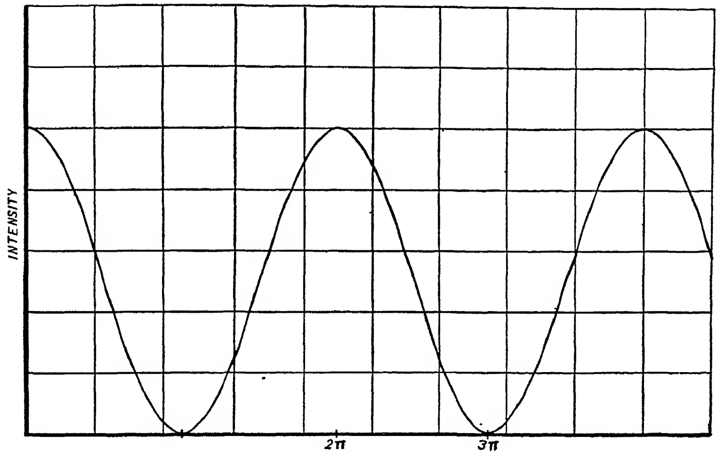


FIG. 102.

The resultant disturbance at P is then given by

$$y = a \sin\left(\frac{2\pi t}{T} - \alpha_1\right) + a \sin\left(\frac{2\pi t}{T} - \alpha_2\right)$$

or
$$y = a\left(\sin\frac{2\pi t}{T}\cos\alpha_1 - \cos\frac{2\pi t}{T}\sin\alpha_1 + \sin\frac{2\pi t}{T}\cos\alpha_2 - \cos\frac{2\pi t}{T}\sin\alpha_2\right)$$

or
$$y = (a\cos\alpha_1 + a\cos\alpha_2)\sin\frac{2\pi t}{T} - (a\sin\alpha_1 + a\sin\alpha_2)\cos\frac{2\pi t}{T}$$

or
$$y = A\cos\theta\sin\frac{2\pi t}{T} - A\sin\theta\cos\frac{2\pi t}{T}, \tag{8.02}$$

if we put $A \cos \theta = a \cos \alpha_1 + a \cos \alpha_2$,

and $A \sin \theta = a \sin \alpha_1 + a \sin \alpha_2$,

from which $$\tan \theta = \frac{\sin \alpha_1 + \sin \alpha_2}{\cos \alpha_1 + \cos \alpha_2},$$

and
$$\begin{aligned} A^2 &= (a \cos \alpha_1 + a \cos \alpha_2)^2 + (a \sin \alpha_1 + a \sin \alpha_2)^2 \\ &= a^2(2 + 2 \cos \alpha_1 \cos \alpha_2 + 2 \sin \alpha_1 \sin \alpha_2) \\ &= 2a^2[1 + \cos (\alpha_1 - \alpha_2)] \\ &= 2a^2(1 + \cos \delta). \end{aligned} \tag{8.01}$$

A and θ, therefore, can readily be found from the original constants a, α_1 and α_2.

Now returning to (8.02), we see that the resultant disturbance is given by

$$y = A \sin \left(\frac{2\pi t}{T} - \theta\right),$$

the equation of an S.H.M. of amplitude A and period T, and differing in phase from either of the original disturbances by an amount which depends on θ and which, therefore, can be found.

As far as amplitude is concerned, we have derived the same expression as by the previous method, and may apply all the results already discussed.

78. Realization of Interference. — The realization of such regions of maximum and minimum intensity depends on our having two such sources as S_1 and S_2, and that is not quite as simple a matter as it may at first thought seem. For example, it would not do to place two sodium flames behind two narrow slits in an opaque screen, and to expect to obtain alternate regions of light and of darkness on a screen near them, for, since the flames are two absolutely independent sources, there can be no regular phase difference between the vibrations of the two slits. At any point P on the screen the phase difference between the two disturbances arriving there is constantly and irregularly changing, and no sustained maxima or minima are possible.

There are several means, however, of realizing the necessary

sources vibrating always with constant phase difference. Of these we shall consider the following.

(*a*) **Double Slit.** — In Chapter I, where reference has already been made to the use of the double slit, it was pointed out that light must first of all spread out from a source slit S, as in Fig. 103. This means that, since well-defined wave-fronts fall on the double slits S_1 and S_2, there will always be a constant phase difference (which may be and often is zero) between the vibrations in the

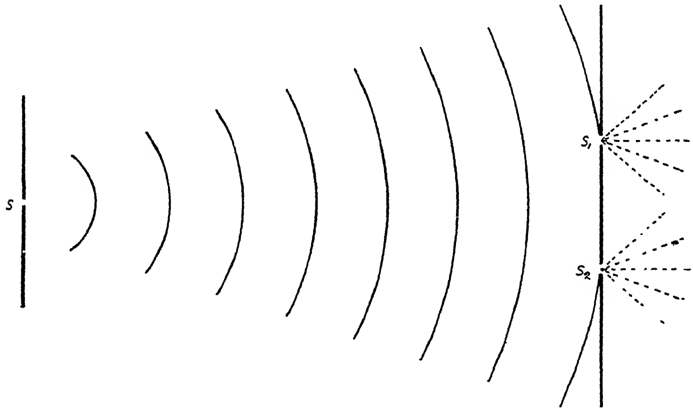


FIG. 103.

separate slits. Interference, therefore, is possible between the two disturbances emerging from the slits, and consequently regions of maximum and minimum intensities (interference fringes) may be observed with this arrangement. Attention has already been directed to Fig. 1, Plate IV, an actual photograph of double slit interference fringes. (See also section 104.)

(*b*) **The Biprism.** — By the use of a double prism such as LM, Fig. 104, with the refracting edge of one portion remote from that of the other, virtual images S_1 and S_2 of a narrow source slit S may be obtained. A screen placed beyond the biprism, therefore, will receive light in the region MN apparently coming

from S_1, in the region AB apparently from S_2. Since each of the apparent sources is the virtual image of the same slit, conditions are again suitable for interference, and in the overlapping

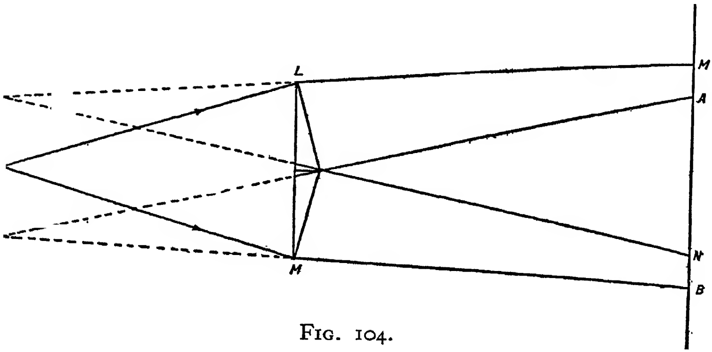


FIG. 104.

portion of MN and AB fringes are observed. Figure 2, Plate IV is a photograph of fringes obtained with the biprism arrangement.

(*c*) **The Fresnel Mirrors.** — The virtual images of a narrow slit S, Fig. 105, formed by reflection at the surfaces of two mirrors M_1 and M_2, inclined at a slight angle, provide still another means of obtaining the necessary conditions for the kind of interference

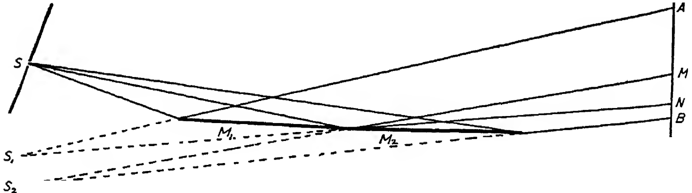


FIG. 105.

we are discussing. No detailed explanation is necessary to show that in the region MN, Fig. 105, light is incident, coming apparently from both the sources S_1 and S_2, and interference fringes are obtained.

(*d*) **Lloyd's Single Mirror.** — By using only one mirror, such as M, Fig. 106, and placing the source slit S near the plane of the reflecting surface, the actual source and its virtual image

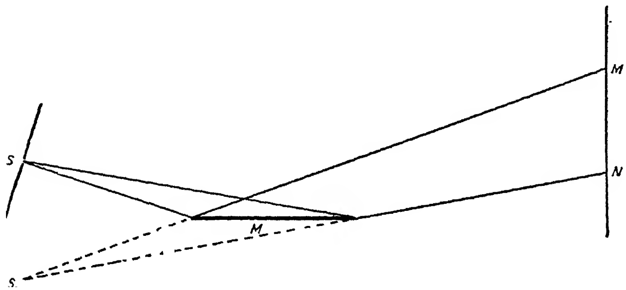


FIG. 106.

S_1 take the place of S_1 and S_2 in the simple diagram of Fig. 105. In this way a fourth method of realizing interference is obtained.

79. Still other methods may be used. We shall proceed, however, to show how simple linear measurements with any of the above arrangements enable us to evaluate the wave-length of the light used.

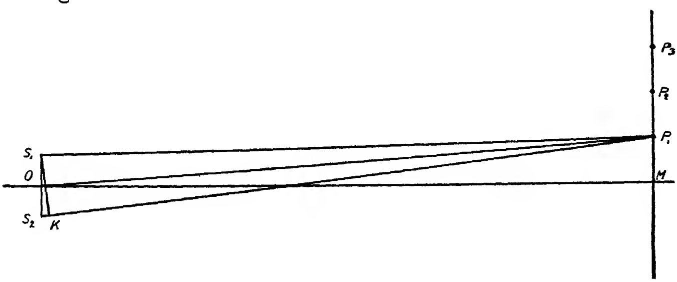


FIG. 107.

Suppose in Fig. 107, that M represents the position of the central maximum (bright fringe), P_1 that of the first to one side, P_2, the second, and so on. We then know that

$$S_2M - S_1M = 0$$
$$S_2P_1 - S_1P_1 = \lambda$$
$$S_2P_2 - S_1P_2 = 2\lambda.$$

Now in an actual arrangement such as the biprism, observation tells us that with a distance of the slit to the screen equal to one or two metres, the distance between bright fringes is invariably less than one millimetre. Angles such as $S_1P_1S_2$, $S_1P_2S_2$, are therefore very small, and the error is slight in putting the sine or the tangent of the angle equal to its measure in radians.

Consequently, if P_1K is taken $= P_1S_1$, so that $S_2K = \lambda$, with little error S_1K may be considered perpendicular to OP_1, where O is the mid-point of S_1S_2.

Also $\angle S_2S_1K = \angle P_1OM$, since OM is $\perp$ to S_1S_2, and OP_1 is $\perp$ to S_1K.

Now in $\triangle P_1OM$, $\tan P_1OM = \dfrac{P_1M}{OM}$

and in $\triangle S_2S_1K$, $\sin S_2S_1K = \dfrac{S_2K}{S_1S_2}$.

since we are dealing with small angles,

$$\frac{S_2K}{S_1S_2} = \frac{P_1M}{OM}$$

or

$$\frac{\lambda}{b} = \frac{x}{D}, \tag{8.03}$$

where b = distance between S_1 and S_2,
D = distance from O to screen,
x = distance from central bright fringe to the next, that is, between any two bright fringes.

Since each of these three quantities b, x, and D may be found by simple linear measurements, we have our first method of calculating wave-lengths. In a certain biprism arrangement, for example, in which $D = 186.03$ cm., $b = 0.654$ cm., $x = 0.0184$ cm., the source of light was an arc lamp screened with a piece of red glass.

The mean wave-length of the light transmitted by this red light, therefore $= \dfrac{0.654 \times 0.0184}{186.03} = 6.46 \times 10^{-5}$ cm.

In the above work we have considered the light approximately monochromatic. If white light is used it is easy to see that the fringes are coloured, except for one or two at the centre, while the number will be very limited. For, since $x = \lambda \dfrac{D}{b}$, the width of the fringes is directly proportional to the wave-length. Thus for $\lambda = .00006$ cm., the second bright fringe from the centre will coincide with the third dark for $\lambda = .00004$ cm. Hence, after the first few fringes, the overlapping will be so great as to give approximately uniform illumination. At the centre of the pattern, however, two or three black and white fringes are observed with coloured ones on either side. In general, for a white light source, only a few fringes can be seen with any arrangement.

80. Wave-Length and Double Slit. — A simple and instructive, although not accurate measurement of the wave-length of

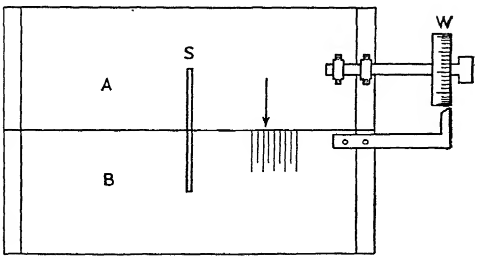


FIG. 108.

light may be made by the use of a double slit and the following method. A source slit S, Fig. 108, is divided into two portions by using two pieces of metal A and B, one of which may be shifted by the wheel W a measured amount with respect to the other. This source may conveniently be illuminated by a 250-watt lamp, with a piece of red glass, R, Fig. 108A, to give approximately

monochromatic light. When an observer focusses a low power telescope T on this source, placed at a distance of two or three metres away, and then inserts a double slit S_1S_2 (see, again, sec-

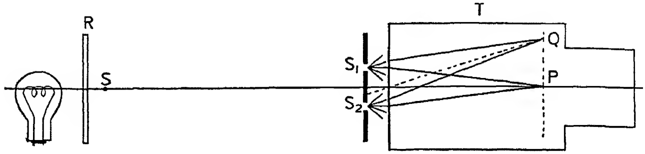


FIG. 108A.

tion 3) in front of the objective, good interference fringes are observed in the telescope because two beams are superimposed at points such as P and Q and they alternately reinforce and counteract each other.

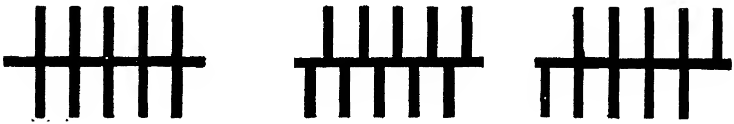


FIG. 109.

By adjusting the position of A until the two portions of the source slit are in line, the fringes appear somewhat as shown in *a*, Fig. 109. On shifting A the fringes at first gradually go out of alignment, as in *b*, Fig. 109, but eventually come into step again,

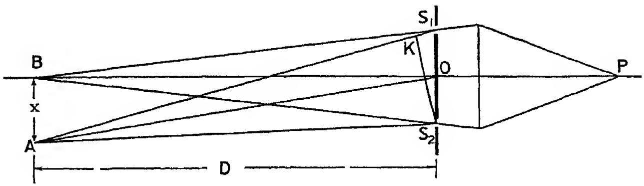


FIG. 109A.

as in *c*, Fig. 109. With continued shift of A, the fringes will of course again go out of alignment and a second time come into step.

Let x = the shift of A, that is, the distance BA in Fig. 109A,

required to bring the fringes into realignment for the first time, as in c, Fig. 109. When this is the case, P, in Fig. 109A, must represent either the central fringe for the fixed portion B of the source slit, or the first fringe to one side of the centre for the movable slit A. This means that the ray paths BS_1P and BS_2P are equal, but that the path AS_1P must be greater than AS_2P by one wave-length. If then AK is taken equal to AS_2, it follows that KS_1 is one wave-length.

If O is the mid-point of S_1S_2, then with little error AO is perpendicular to S_2K, and

$$\angle KS_2S_1 = \angle AOB.$$

Putting this angle $= \phi$, we can write

from $\triangle AOB$, $$\tan\phi = \frac{AB}{BO} = \frac{x}{D},$$

and from $\triangle KS_2S_1$, $$\sin\phi = \frac{KS_1}{S_1S_2} = \frac{\lambda}{b}, \text{ where } b = S_1S_2.$$

Since ϕ is small, $\tan\phi = \sin\phi$, and we have

$$\frac{}{b} = \frac{}{D},$$

or $$\lambda = \frac{xb}{D}.$$

Here x is the shift of the movable slit A corresponding to a shift of one fringe; b is the width from centre to centre of the two apertures of the double slit; and D is the distance of the source slit to the double slit. All of these are easily measured quantities.

The following readings were taken by students when using the above arrangement:

Mean value of b for slit I = 0.0341 cm.
Mean value of x for slit I = 0.586 cm.
when distance D = 317.9 cm.
Hence wave-length = 6.28×10^{-5} cm.

Mean value of b for slit II = 0.0875 cm.
Mean value of x for slit II = 0.227 cm.
when distance D = 317.6 cm.
Hence wave-length = 6.25×10^{-5} cm.

81. One or two applications of interference will next be given, the first of which is illustrated by the following problem.

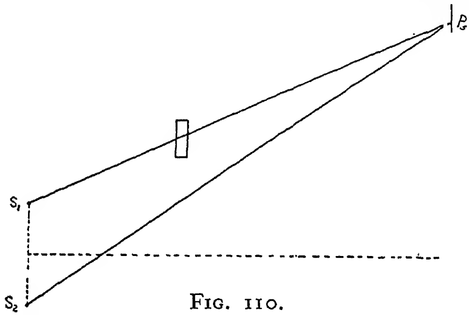


FIG. 110.

When a thin flake of glass, with index of refraction = 1.5, *is introduced into the path of one of the interfering beams in such an arrangement as the biprism, the position of the central bright fringe becomes that normally occupied by the fifth. If the wave-length used is* 6×10^{-5} *cm., find the thickness of the glass.*

In the normal arrangement (that is, without the glass flake), if P_5, Fig. 110, is the position of the fifth bright fringe, then

$$S_2P_5 - S_1P_5 = 5\lambda, \text{ where } \lambda = \text{wave-length in air.}$$

But with the glass flake present, since P_5 is the position of the central fringe, the optical path $S_2P_5 - S_1P_5 = 0$.

Therefore, if t cm. = required thickness of glass, *replacing* t *cm. of air by* t *cm. of glass must lengthen the optical path* S_1P_5 *by* 5λ.

But $\quad t$ cm. of glass = $1.5\,t$ cm. of air,

$\therefore \quad$ the path *change* = $(1.5 - 1)t$ cm. of air.

Hence

$$(1.5 - 1)t = 5\lambda$$
$$= 5 \times 6 \times 10^{-5},$$

from which

$$t = 6 \times 10^{-4} \text{ cm.}$$

Hence, given the index of a thin piece of glass, and the wave-length of the light, we can find its thickness. Conversely, given the thickness and the wave-length, the index can be found.

82. Rayleigh's Refractometer.* — Some interesting applications of interference are provided by the use of this instrument,

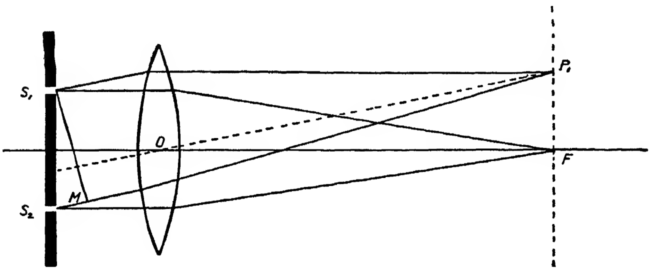


FIG. 111.

which, in principle, is just a double slit placed before the objective of a telescope. When a distant narrow source, a star for example, is viewed through such an arrangement, fringes are observed in the focal plane of the objective.

Suppose the finite size of the star can be neglected, and that the parallel rays from it are incident on the telescope in a direction parallel to the axis of the instrument. Then, F, Fig. 111, the principal focus of the objective, is also the position of the central bright fringe. If P_1 is the position of the first bright fringe to one side of the central, then, since $\angle\ P_1OF = \angle\ S_2S_1M$, the angular separation of two bright fringes is measured by $\frac{S_2M}{S_2S_1}$ or $\frac{\lambda}{b}$, since $S_2M = \lambda$ and $S_1S_2 = b$. (If the student has difficulty in seeing this, he should re-read p. 9, Chapter I, bearing in mind that in this arrangement the focal plane of the objective replaces the retina of the eye in Fig. 7.)

* See also sec. 104.

It follows at once that, if the distance between the two slits is adjustable, then the width of the fringes may be varied, being in-

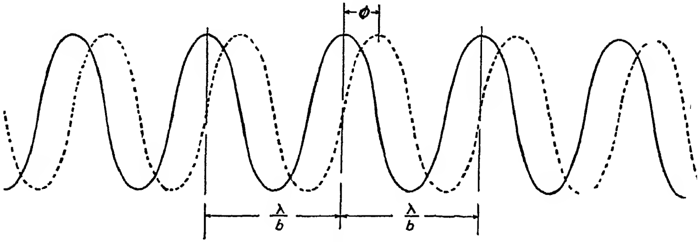


FIG. 112.

versely proportional to b. The position of the central fringe, however, remains unaltered, since it is determined by the direction the incident light makes with the axis of the telescope, and, indeed, coincides with the image of the star formed by the telescope in the absence of the double slit.

Suppose, now, that a double slit telescope is directed toward two stars close together. Then, for each star, there is a system of fringes of angular width $\frac{\lambda}{b}$. The positions of the central fringes, however, are not the same, being separated by an angular distance ϕ, equal to the angle the two stars subtend at the telescope. (Recall again that the central fringe for each system coincides with the position of the image of the corresponding star in the absence of the double slit.) We then have two superimposed sets of fringes, somewhat as shown in Fig. 112. Now let the magnitude of b be increased. The angular width of the fringes then decreases, and, for a particular value of b, the maximum of one will fall on a minimum of the other, as in Fig. 113. The resultant illumination will then be uniform, no fringes being observed. When this is the case, we have

$$\phi = \frac{\lambda}{2b}.$$

Consequently such an arrangement provides us with a means of

measuring the angular separation of two close stars, from the observation of the value of b for which the fringes disappear.

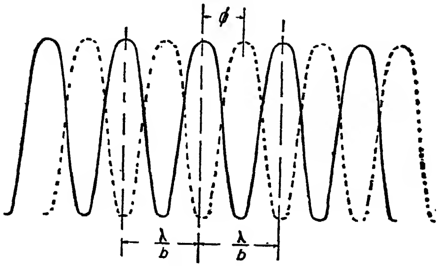


FIG. 113.

This work is based on the assumption that the dimensions of the star itself may be neglected. Now, although the angular width of a star is small, it is finite. It follows, therefore, that if we subdivide the surface of the star into a number of elemental strips, and consider each of these as a source of light, for each there will be a corresponding set of fringes with maxima in slightly different positions. In the case of a single star of finite width, therefore, we have a *slight* overlapping of a large number of systems of fringes. With small values of b, the fringes are so wide and the overlapping so slight, that no noticeable blurring of the fringes is observed. In that case the star may be treated as a point. That this is not always so, however, has been shown by a wonderful piece of experimental work done by Prof. Michelson of the University of Chicago. Using the star α Orionis, and very large values of b, he showed not only that the overlapping could not be neglected, but also that it could be used as a means of measuring the angular diameter of the star. Although the mathematical analysis dealing with the case of a single source of finite width is beyond the scope of this book, the student should not have difficulty in seeing generally, that if the overlapping is sufficient the fringes disappear altogether, and that the condition for such disappearance depends both on b and on the angular

diameter of the star. The analysis shows that the fringes disappear for a value of b given by

$$\text{angular diameter} \quad 1.22\frac{\lambda}{b}.$$

In the case of α Orionis, Michelson found that when b = 121 inches, the fringes disappeared, although they were sharp for other

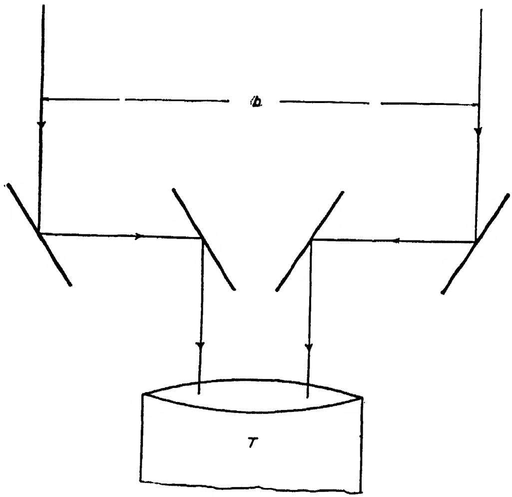


FIG. 114.

smaller stars. Taking λ = 5750 A, this gives 0.047 sec. for the angular diameter of this star. To make this amazing measurement great skill and ingenuity was required. The large value of b was obtained by having a framework carefully constructed and attached to the telescope. By means of a system of adjustable mirrors carried by the framework, light was reflected somewhat as shown in Fig. 114.

83. Index of Refraction of Gas. — When a narrow slit is placed in the focal plane of a good lens, conditions are essentially the same as if the slit were at infinity. For example, if a spectrom-

eter is adjusted in the usual way and a double slit placed on the table (or immediately in front of the objective of the telescope) with its face normal to the beam emerging from the collimator, excellent interference fringes are observed in the telescope. Since, with this arrangement, the *angular* width $\frac{\lambda}{b}$ of the fringes can be measured, although not accurately, an estimate of the wave-length of the light used may readily be obtained.

In one form of the Rayleigh's Refractometer the arrangement of Fig. 115 is used. This form is adapted for use in measuring the index of refraction of air or other gas by the principle of interference. Suppose two tubes, A and B, with good glass windows closing the ends, are placed side by side in the path of the parallel beam between the collimating lens L and the double slit, so that light incident on S_1 traverses one tube, that on S_2 the

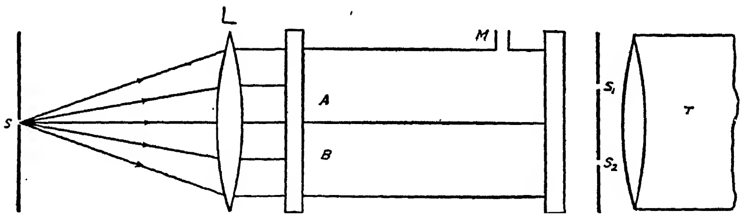


FIG. 115.

other. Then, if, by means of an outlet M attached to tube A, the density of the air (or other gas) in this tube is altered, the optical path of one of the interfering beams is altered and the position of the fringes will change. Technically it is said that there is a *shift* of the fringes, for, on viewing them, one observes a lateral motion as the position of a maximum (or minimum) changes. By observing the fringe shift for a given alteration in density (or pressure), it is possible to calculate the index of the gas in this tube. Again the principle of the method may be illustrated by the following example, which will be worked in three different ways.

Problem: *When all the air is removed from the tube A, Fig. 115,*

a shift of 150 fringes is observed. If the wave-length in air (at room temperature and the initial pressure) is 0.00004 cm. and the tube is 20 cm. long, find the index of refraction of the air.

Method I.

In this problem it is necessary to distinguish carefully between air and vacuum. Thus, if $n =$ the required index, then, the wave-length in vacuo $= 0.00004n$ cm.

Since the fringe shift $= 150$, that is, since at a given place a bright fringe changes from bright to dark and back again to bright 150 times,

$\therefore$ the path change brought about by removing the air in tube A

$$= 150 \text{ wave-lengths}$$
$$= 150 \times 0.00004 \text{ cm. air}$$
$$= 150 \times 0.00004n \text{ cm. vacuum.}$$

But, since the tube is 20 cm. long,

the initial path in A $= 20$ cm. air
$= 20n$ cm. vacuum,

the final path in A $= 20$ cm. vacuum.

$\therefore$ the path change also $= 20(n - 1)$ cm. vacuum.

Hence, $20(n - 1) = 150 \times 0.00004n,$

from which $n = \dfrac{20.000}{19.994} = 1.0003.$

Method II.

The number of wave-lengths in 20 cm. air $= \dfrac{20}{0.00004}$.

the number of wave-lengths in 20 cm. vacuum $= \dfrac{20}{0.00004n}$,

$\therefore$ the *change* in the number when the air is removed

$$\frac{20}{0.00004} \qquad \frac{20}{0.00004n}$$

But every time there is a change in the number by 1, there is a fringe shift of 1.

$$\therefore \quad \frac{20}{0.00004} - \frac{20}{0.00004n} = 150,$$

from which $$n = 1.0003.$$

Method III.

Let
$$V = \text{vel. of light in vacuo},$$
$$V_a = \text{velocity in air, so that}$$
$$V = n \cdot V_a.$$

Then, the time for light to travel 20 cm. in air $= \frac{20}{V_a} \quad \frac{20n}{V}$,

the time for light to travel 20 cm. in vacuum $= 20$

$$\therefore \quad \text{difference in times} = \frac{20}{V}(n - 1).$$

But the time difference is equal to that corresponding to a path difference of 150 wave-lengths, and therefore $= 150T$, where T is the periodic time.

$$\frac{20}{V}(n - 1) = 150T.$$

But
$$\lambda \text{ vacuum} = VT$$
$$\frac{20(n - 1)}{V} = \frac{150 \cdot \lambda \text{ vacuum}}{V}$$
$$20(n - 1) = 150 \times 0.00004n,$$

from which $$n = 1.0003.$$

As a matter of fact a fringe shift as great as 150 could not be obtained with such an arrangement. Actually, the fringe shift for a small pressure change is observed. In that case the index of air at 0°C and 760 mm. pressure is found from the following relation:

$$n - 1 = f \cdot \lambda \cdot \frac{1}{273} \quad \frac{76}{p_1 - p_2} \quad \frac{\ }{L},$$

where
$$f = \text{fringe shift},$$
$$T = \text{absolute temperature},$$
$$p_1 = \text{initial pressure},$$
$$p_2 = \text{final pressure},$$
$$L = \text{length of tube path}.$$

CHAPTER IX

INTERFERENCE (Continued)

Thin Films and Plane Parallel Surfaces

84. In the work of the previous chapter, interference fringes were obtained by superimposing two disturbances, each of which originally came from the same source, although by two different paths. In all the cases considered it was essential that a narrow source be used, for otherwise, as briefly indicated in the case of a star of finite magnitude, different parts of a broad source give rise to maxima in different places, and uniform illumination results. It is the purpose of this chapter to discuss in detail other ways of obtaining interference fringes, in all of which a *broad* source is necessary. The brilliant colours seen when light is reflected from a thin layer of oil on a wet pavement, and the entrancing colours every child has seen when blowing soap bubbles, are familiar examples of this kind of interference.

Suppose a soap film, formed at the mouth of a beaker and strongly illuminated with white light, is viewed directly with the naked eye, or, alternately, that an image of the film is cast on a screen by a projecting lens. If the film is in a vertical position, a series of beautifully coloured horizontal bands are observed. If the illuminating source is made approximately monochromatic by placing a piece of red glass in the path of the light, a series of alternately red and black horizontal fringes are obtained.

In seeking the explanation of this experiment, it is necessary to remember that a soap film is in reality a thin layer of water enclosed between two surfaces, in this case, plane, in the case of a soap bubble, spherical. The film at the mouth of the beaker is wedge-shaped, gradually growing thinner and thinner at the top as the water falls due to gravity, and as it evaporates. Now when light falls on such a transparent film, part of the beam is reflected

at the first surface, part is transmitted, of which a portion is reflected at the second surface, this portion eventually emerging from the first face. In the reflected light, therefore, there are two disturbances, both of which originally came from the incident beam, but one of which travelled a distance, approximately equal to twice the thickness of the film, greater than the other. Interference is then possible. Because of the changing thickness of the soap film, with monochromatic light regions of maxima and

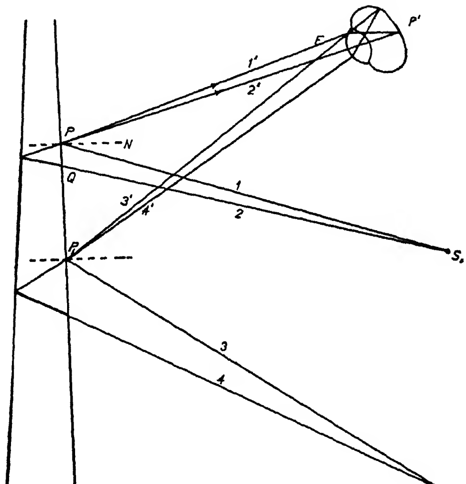


FIG. 116.

minima alternate and fringes are observed. If white light is used, then for any given thickness, the resultant intensity will depend on the wave-length. It might be a maximum for one colour, a minimum for another, so that, when all colours are present, certain portions of the original beam are either feebly present or absent altogether, and at all places the resultant is coloured.

To discuss this type of interference more in detail, it is convenient to make use of individual rays of light. Suppose an eye is focussed on the small region about P, Fig. 116, on the first face

of a thin film illuminated with monochromatic light. The eye can then receive light by means of rays 1′ and 1 from a *small source* S_1 so placed that, if PN is normal to the film, $\angle EPN = \angle NPS_1$. Another ray 2 may leave the source S_1, strike the first surface at a point Q near P, be refracted, then reflected at the second surface, and eventually emerge at P, travelling to the eye as ray 2′. The image of P on the retina, P', is formed, therefore, by means of two disturbances, both coming originally from S but travelling by

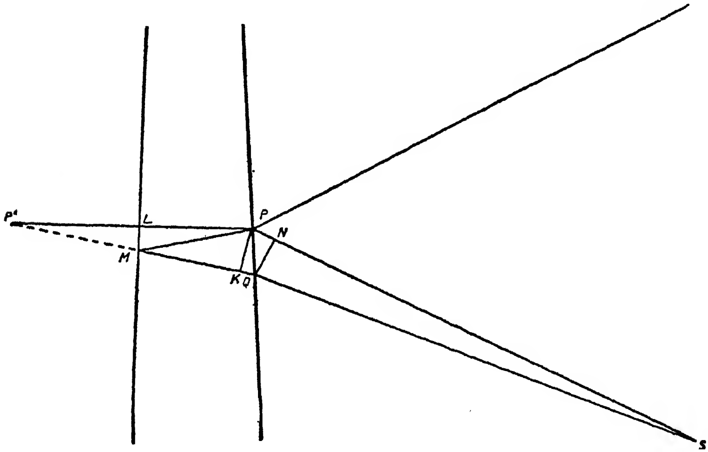


FIG 117.

different paths. Consequently the intensity at P' may vary all the way from a maximum to a minimum, depending on the magnitude of the path difference.

If a broad source of light is used, with the eye in the same position, the images of other places such as P_1 are formed by means of other pairs of rays, such as 3′ and 4′ which come originally from the small portion S_2 of the source. It follows, therefore, that if the thickness of the film gradually increases, the path difference between such pairs of interfering rays will do so also and consequently the face of the film will be alternately light and dark. That this is the case has already been pointed out in the soap

film experiment. It may also readily be shown by forming a thin air wedge between two pieces of plate glass, touching at one end, slightly separated at the other, and viewing the reflected light from a yellow sodium flame.

Pairs of rays such as 1′ and 2′, 3′ and 4′ are called *congruent rays*. For any such pair we wish now to calculate the path difference. To do so, consider Fig. 117, bearing in mind always that, for the purpose of making the argument clear, in both this figure and in Fig. 116, the thickness of the film and the distance PQ are grossly exaggerated. Thus, in Fig. 117, PQ and LP are actually extremely small quantities compared with distances such as SP or SQ.

The angle PSQ therefore is of small magnitude, and so if we draw $QN \perp$ to SP, SQ and SN are equal, to the first order of small quantities.

$\therefore$ the difference between the paths $SQMP$ and SP

$$= QM \text{ (in film)} + MP \text{ (in film)} - NP \text{ (in air)}$$
$$= n \cdot QM + n \cdot MP - NP \text{ cm. of air,}$$

if $n =$ index of the film material.

Now draw $PK \perp$ to MKQ.

Then since $\angle NQP =$ angle of incidence at P,

and $\angle KPQ =$ angle of refraction at Q,

$$PN = n \cdot KQ.$$

$\therefore$ required path difference

$$= n \cdot QK + n \cdot KM + n \cdot MP - PN$$
$$= n(KM + MP).$$

Produce KM to P', making $MP' = MP$, and join PP'.
Then, the path difference

$$= n(KM + MP)$$
$$= n(KM + MP')$$
$$= n \cdot KP'$$
$$= n \cdot PP' \cos PP'M.$$

By elementary geometry, it is easily seen that PP' must be perpendicular to the second face, and that $LP' = LP$.

$\therefore \quad PP' = 2t$, where t is the thickness of the film at P.

Also
$$\angle PP'M = \angle P'PM$$
$$= (r + \theta),$$

where r is the angle the internal ray makes with the normal to the first surface at P, and θ is the angle of the wedge formed by the film.

Therefore, the path difference

$$= 2nt \cos (r + \theta). \qquad (9.01)$$

Now we have to do only with films whose surfaces are almost parallel, that is, with those for which θ is such a small quantity that, with negligible error, we can put $\cos (r + \theta) = \cos r$.*

We have then finally, path difference

$$= 2nt \cos r. \qquad (9.02)$$

Several cases to which this important formula for the path difference is applicable, are of interest.

85. Fringes of Equal Thickness. — If an observer views a piece of film a few centimeters in diameter at a distance of 25 cm. or more, the variation in $\cos r$ in going from point to point on the film may be small compared with the variation in t. In that case the variation in t is largely responsible for the change from maximum to minimum, and the fringes are sometimes called *fringes of equal thickness*. This means, of course, that a bright fringe is the locus of points corresponding to places of equal thickness. The wedge-shaped soap film fringes, for example, belong to this class, as well as those in the Newton's Rings arrangement. (See section 87.)

* The student who finds difficulty in steps of this sort will do well to take a concrete case. Thus, if a thin air wedge is formed by separating one end of two plate glass pieces, 10 cm. long, by a thin piece of glass 0.04 mm. thick, the angle θ has the value .0004 radian or a little over 1 min. Taking a case where $r = 30°$, we find $\cos 30° = .8660$, $\cos 30°1' = .8661$. Hence, in such a case, the error in putting $\cos (r + \theta) = \cos r$ is less than 0.015%.

86. In thin film fringes if $t = 0$, the path difference is zero, and one might expect the resultant phase difference to be zero also. Now just before the soap film breaks, one can easily notice that at the top (where it is thinnest) it is dark, not bright. In other words, although the path difference is zero (for the two surfaces come together just at breaking), the two disturbances are out of step, or the phase difference is 180°. The reason for this is found in the fact that the conditions of reflection at the two surfaces are not the same. At the first face, reflection takes place at a surface

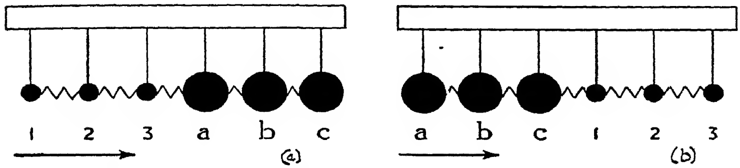


Fig. 117A.

separating a less dense medium (air) from a more dense (water), whereas at the second face the opposite is the case. In the former case there is a sudden change of phase of π radians in the act of reflection, but no such change in the latter case. The student may find some help in realizing this fact by recalling the reflection of sound at the end of a closed and an open organ pipe. In the former case, a node always exists at the closed end, that is, air particles adjoining the closed end are always at rest. Consequently, there must always be a phase difference of 180° between the direct and the reflected waves, for this is the only condition which would keep these particles at rest. In the open organ pipe the end particles can move more freely than those inside the pipe and no such change of phase occurs when reflection takes place.

The difference between the two cases may be explained by considering a row of particles in elastic contact, as represented in Fig. 117A, along which a longitudinal wave disturbance is travelling. We distinguish two cases; (1) when the disturbance, travelling along particles of small mass, such as 1, 2 and 3 in (*a*), strikes a region of heavy particles, such as *a*, *b* and *c*; (2) when the dis-

turbance, initially along heavy particles, encounters those of small mass, as in (*b*). In the first case when the disturbance reaches particle 1, displacing it in the direction of the arrow, this particle then pushes against 2, 2 against 3, and 3 against the heavy particle *a*. Because the mass of *a* is much greater than 3, particle 3 bounces back, that is, its motion is suddenly changed into one in the opposite direction. Such a sudden change in motion is equivalent to a phase difference of half a vibration, that is, of 180°. Due to this sudden reversal, particle 3 now pushes back against 2, and a reflected wave travels backwards. This case corresponds to the reflection of light incident on glass after traversing a less dense medium like air.

In the second case, where the disturbance, after traversing the heavy particles, encounters those of small mass, as in (*b*), the end particle *c* is able to move forward more readily than those preceding it because its motion is less hindered by the particle of small mass adjoining it. Since particle *c* moves forward in the same direction as before its impact with 1, there is no change of phase in its motion, although a reflected disturbance again travels backwards. This case corresponds to the reflection of a beam of light, which, after traversing a medium like glass, encounters a less dense medium like air.

87. Newton's Rings. — If fringes of equal thickness are viewed normally, then since the cosine of a small angle = 1, the path difference = $2t$. A good example is found in the Newton's Rings arrangement as represented in Fig. 118. In this figure L represents a plano-convex lens whose curved face rests on the plane face of a glass block P. A thin film of air whose thickness increases radially from the point of contact is then formed between the curved and the plane glass surfaces. Light from a sodium flame F, collected by the lens C, falls on a thin inclined transparent plate R and is reflected downwards to the film. The light reflected back again from the faces of the film is viewed by a low power travelling microscope M focussed on the portion of the film about the point of contact. Interference fringes are then observed for

exactly the same reason as has been explained for the soap film. In this case, since the path difference at any point $= 2t$ and since regions of equal thickness are obviously circular, the fringes observed are circular. Because of the different conditions of reflection at the two faces, the centre of the fringe system is dark. A reproduction of a photograph of such fringes is given in Fig. 1, Plate VI.

Such an arrangement provides us with another means of evaluating the wave-length of light by measuring, (1) the diameter

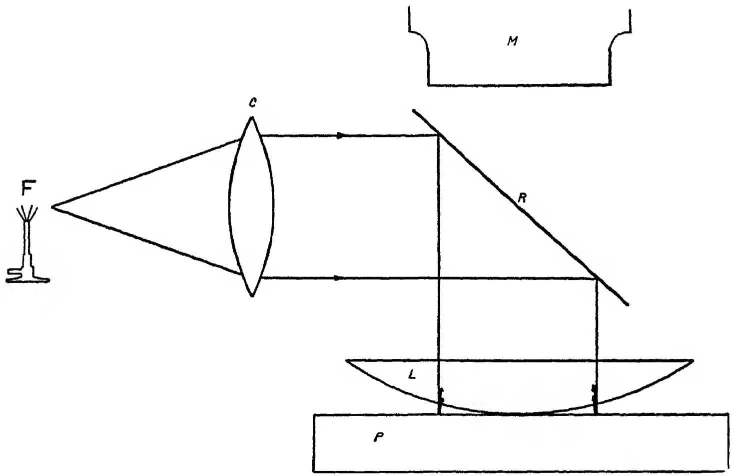


FIG. 118.

of any ring, (2) the radius of curvature of the upper curved surface. To find the necessary expression for the calculation, consider the nth dark ring. At the centre (which again is dark), the path difference $= 0$,

$\therefore$ at the 1st dark ring, the path difference $= \lambda$,
at the 2nd dark ring, the path difference $= 2\lambda$.

and generally at the nth dark ring, the path difference $= n\lambda$.

But the path difference anywhere $= 2t$,

$$\therefore \; n\lambda = 2t. \tag{9.03}$$

Now, if DN Fig. 119, is the radius of the nth dark ring, then

$$AN = t = OD.$$

But, by elementary geometry, as already proved in section 31, Chapter III,

$2R \cdot OD = DN^2$, where R is the radius of curvature of the curved face.

$$\therefore \quad OD = \frac{DN^2}{2R} = \frac{D_n^2}{8R},$$

where D_n = diameter of the nth dark ring.

Finally, we have

$$2t = \frac{D_n^2}{4R},$$

or

$$\lambda = \frac{D_n^2}{4nR}. \tag{9.04}$$

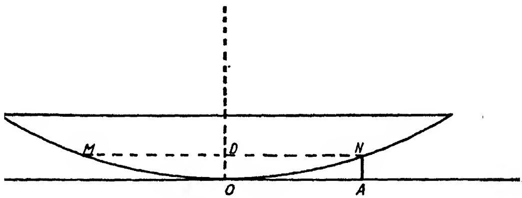


FIG. 119.

Now it is a simple matter to measure D_n by means of the travelling microscope, and R can easily be found by a spherometer, hence we have another, although again, not a particularly accurate method, of evaluating λ.

If the nth bright ring is used, it is left as an exercise for the student to prove that

$$\lambda = \frac{D_n^2}{2R(2n - 1)} \tag{9.05}$$

Since examples of the kind of interference we have been discussing are common in nature, it is not surprising that Newton and some of his contemporaries were familiar with the colours

of thin plates, and the rings to which the great Newton's name is given. To Hooke, who in 1664 published a book one chapter of which dealt with this phenomenon, we owe the statement — "In general wheresoever you meet with a transparent body thin enough, that is terminated by reflecting bodies of differing refractions from it, there will be a production of these pleasing and lovely colours." Indeed, while Newton explained the existence of fringes by applying ideas analogous to those involved in his theory of fits of easy reflection and easy transmission, in such a way that for certain thicknesses rays were transmitted, for alternate thicknesses reflected, Hooke's explanation almost anticipated the modern view, and that more than one hundred years before the pioneer work of Young on interference. According to Hooke light was reflected at both the front and the rear surfaces of a thin film, coloured phenomena resulting from the delay the latter beam suffers in crossing and re-crossing the film.

88. Fringes of Equal Inclination. — Suppose we have a thin film whose faces are parallel. Since t everywhere is constant, r is the only variable, and fringes can be observed only if the variation in $2nt \cos r$ over the face of the film is great enough to give path differences of a few wave-lengths. In such a case, if a perpendicular is dropped from the eye to the film it is viewing, it is not difficult to see that the fringes obtained will be circular, with the foot of the perpendicular as centre.

If t is large, that is, very many times the wave-length of light, fringes such as we have so far been discussing in this chapter cannot be obtained. To understand why this is so, it is first of all necessary to recall that the pupil of the eye has a finite aperture, of the order of 3 mm. That being the case, the eye can form an image of P, Fig. 120, by means of other pairs of congruent rays, for each of which the value of r is slightly different. Thus, for one pair the path difference is $2nt \cos r$, for another pair $2nt \cos (r + \triangle r)$, where $\triangle r$ is a small fraction of r. Now if t is not more than a few wave-lengths, the difference between these two values is a small fraction of a wave-length. If, there-

fore, P is bright as observed by one pair, it will be so also for other pairs.

If t is large, a few millimetres or a centimetre or two, then the difference between $2nt \cos r$ and $2nt \cos (r + \triangle r)$ becomes sufficiently great that one pair of congruent rays may give rise to a maximum, another to a minimum, and so on. In such a case fringes cannot be obtained.

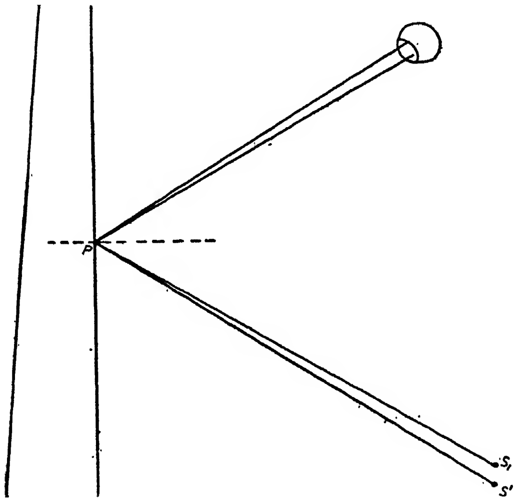


FIG. 120.

If, however, the surfaces are parallel and optically plane (that is, if the actual surface deviates from a geometrical plane by a small fraction of a wave-length), then fringes of another kind can be obtained even when the surfaces are separated many centimetres. This type of interference can be observed by the naked eye, or more usually by means of a telescope, provided each is focussed for infinity, that is, so that parallel rays are brought together at a point on the retina, or in the focal plane of the telescope objective.

Suppose AB, CD, Fig. 121, represent two such surfaces, plane and parallel to each other. Then, since at each surface, a partial

reflection and a partial transmission is possible, an incident ray such as 1 can give rise to a reflected series of parallel rays a, b, c, d, e . . . as well as a transmitted series a', b', c', d', e' . . . . If such a bundle falls on the objective of a telescope, a point image P is formed in its focal plane. Now between every two of these parallel rays there is the same path difference, which, by a method similar to that given above, may readily be shown to be equal to $2nt \cos r$, and which, therefore, since n and t are constants, varies only with r. It follows that the resultant in-

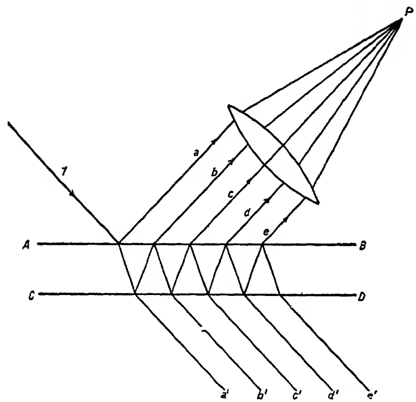


FIG. 121.

tensity of the image at P varies with r, alternating from a maximum when $2nt \cos r$ is equal to a whole number of wave-lengths, to a minimum when it is equal to an odd number of half wave-lengths. If then a beam incident on the face AB contains many rays striking the surface at different angles, and if either the reflected or the transmitted light is observed by a telescope focussed for infinity, then a series of fringes is observed. Such fringes are called *fringes of equal inclination.*

To make the matter more concrete, a brief reference will be made to a simple instrument called the Fabry and Perot étalon,

by means of which fringes of this kind are obtained using transmitted light. This instrument consists essentially of two glass plates, the inner adjacent surfaces of which, AB, CD, Fig. 122, are optically plane, are accurately parallel, are "half" silvered, and separated by an air layer. As explained in connection with Fig. 121, if *approximately* parallel light is incident on this arrangement, then, due to multiple reflections and transmissions, we have

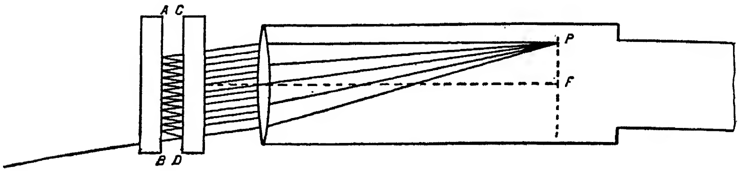


FIG. 122.

many emergent parallel bundles of rays, only one of which is shown in Fig. 122. Since there is an infinite number of parallel bundles all of which make the same angle with the normal to the silvered surfaces, then, if P is a maximum (bright fringe), it follows that all points on the circle passing through P with F as centre, are also bright. In other words we have circular fringes similar to those reproduced in Fig. 4, Plate IV.

The fringes obtained by such an arrangement differ from those obtained by a biprism or a double slit, in respect to the distribution of intensity when one goes from one bright fringe to the next. A comparison of Fig. 2, Plate IV (biprism fringes), or of Fig. 1,

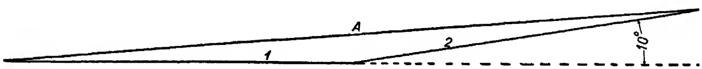


FIG. 123.

Plate IV (double slit), with Fig. 4, Plate IV (Fabry and Perot), will show that in the latter case the maxima (bright fringes) are much narrower than in the others. The reason for this is found in the fact that there are several interfering beams. This will be evident if the student will apply the method of section 15, Chapter II, to find

PLATE IV

FIG. 1. Double Slit Interference Fringes.

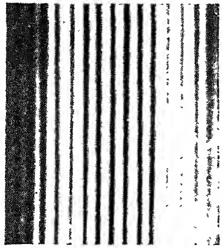


FIG. 2. Biprism Fringes.

FIG. 3. Lummer Plate Fringes.

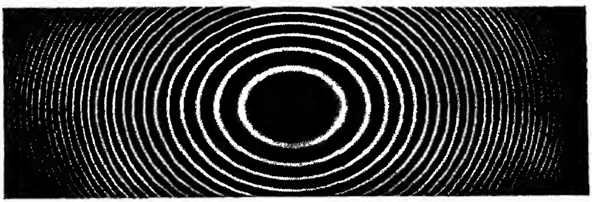


FIG. 4. Fringes obtained with Fabry and Perot Étalon.

FIG. 5. Interference fringes obtained with quartz wedge and polarized light.

the resultant intensity, (1) of two disturbances differing in phase by 10°; (2) of eight, between every two of which there is a phase difference of 10°. It will be found, by making diagrams similar to Fig. 123 and Fig. 124, that in the former case the intensity is

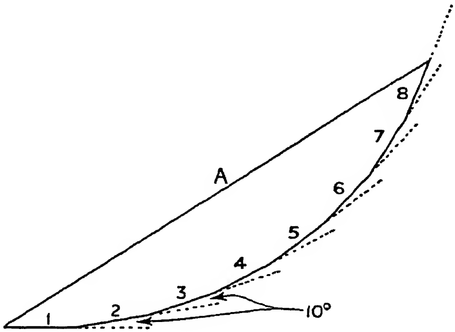


FIG. 124.

more than 99% of the maximum, while in the latter it has dropped to about 86% of the maximum. This number has been obtained on the assumption that all the beams are of equal amplitude,

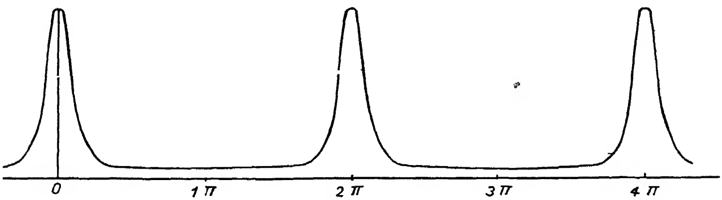


FIG. 125.

whereas as a matter of fact the rays *a′*, *b′*, *c′*, *d′*, *e′*, Fig. 121, steadily decrease in intensity. The general result, however, is not altered when this is taken into consideration, for the complete analytical solution shows that the curve giving the distribution of intensity for Fabry and Perot fringes is somewhat as illustrated in Fig. 125. If this is compared with the corresponding figure for

two interfering beams, Fig. 102, it will at once be evident that many interfering rays give rise to narrow maxima.

Because of this fact, such an arrangement has several useful applications, two of which will be briefly considered.

89. Comparison of Two Wave-Lengths. — An important use to which the Fabry and Perot étalon has been put is in the accurate comparison of wave-lengths. This is possible because the value of the radius of any fringe is a function of the wave-length. If now two wave-lengths in a source are utilized, there are two systems of rings whose radii may be accurately measured. If one of these wave-lengths is known to a high order of accuracy, then the wave-length of the other may be found to the same degree of accuracy by using a simple relation involving radii of corresponding rings.

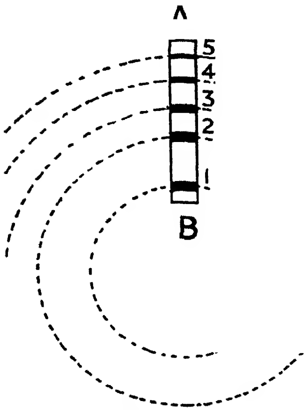


Fig. 125A.

90. Structure of a Single Spectral Line. — An important use to which the Fabry-Perot étalon is often put is the analysis of a single spectral line. By placing the interferometer in the path of the beam of light between the collimating lens and the prism of a spectrograph (see section 58) we obtain not only the usual spectrum of the source but also part of a system of circular fringes crossing each spectral line. The appearance is then somewhat as shown by the heavy lines in Fig. 125A, where AB represents the ordinary slit image of the spectral line, now, however, crossed with short interference fringes (which in reality are part of a circular system indicated by the dotted lines). Reproductions of actual photographs taken in this way are given in Figs. 3 and 4, Plate VII.

Now when the light corresponding to a spectral line is examined after this fashion it is generally found that the pattern obtained is not just a series of single fringes as in Fig. 125A or Fig. 3 (*a*), Plate VII. More often the pattern consists of a group of two or more fringes regularly repeated. An excellent example of this is shown in Fig. 3 (*b*), Plate VII, for which, together with Fig. 3 (*a*), the author is indebted to Dr. E. Gwynne Jones of the Imperial College, London, England. Fig. 3 (*b*) clearly shows that what in the ordinary spectrograph appears to be a single spectral line, (in this case, one due to bismuth, of wave-length 6600 A), in reality has three components. The reason they appear as a single line in the ordinary spectrograph is due to the fact that their wave-lengths are too close together to be separated by the dispersion of prisms. Actually in this case, the components differ by 0.65 angstrom and 0.53 angstrom.

In Fig. 4, Plate VII, another example of this type of analysis is given. In Fig. 4 (*a*), the interference pattern, that of the red hydrogen line H^1_α, of mean wave-length 6562.79, shows that this line has at least two components and actual measurement gives their separation as 0.14 angstrom. Fig. 4 (*b*) is the interference pattern obtained with exactly the same optical arrangement as Fig. 4 (*a*), but with a luminous source containing a mixture of ordinary and of heavy hydrogen. In section 167A it will be explained that the red line emitted by heavy hydrogen, that is, H^2_α differs from the corresponding wave-length for ordinary hydrogen by 1.79 angstroms. The additional set of fringes which appears in Fig. 4 (*b*) is then due to this heavy hydrogen line. Obviously this line has the same double structure as H^1_α.

Of the many spectral lines which have been examined in this way, there are extremely few which do not show a fine structure of some kind. The red cadmium line 6438 which is used as a standard in the accurate measurement of wave-lengths, is about the nearest approach to monochromatic light which has been found.

Interferometers of high resolving power are also used for the examination of the Zeeman effect (see section 75) when the component lines differ from each other by a small fraction of an angstrom.

For example, in Fig. 4, Plate III, the two components there shown differ by only 0.07 of an angstrom.

91. Lummer Plate. — Parallel bundles of interfering rays may be obtained by means of another comparatively simple device called the Lummer-Gehrcke plate, or sometimes the Lummer plate. While the essential part of the instrument is just a long narrow glass (or quartz) plate, its effectiveness lies in the fact that the sides *AB* and *CD*, Fig. 126, are optically plane and parallel. When, therefore, by means of a small reflecting prism *P* at one

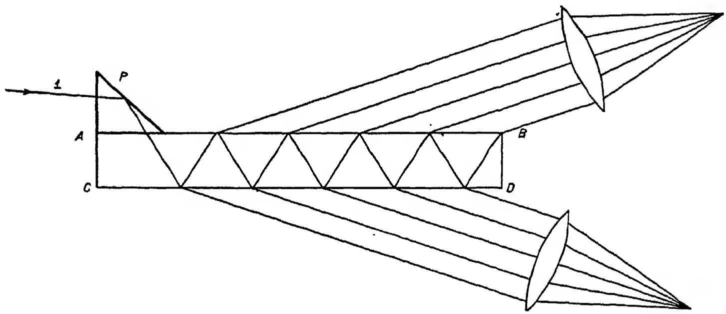


FIG. 126.

end, a ray 1 is introduced into the plate, multiple reflections give rise to bundles of parallel rays on either side of the plate, between every two of which there is the same path difference. For the same reason, and with the same general optical arrangement as in the Fabry and Perot étalon, interference fringes, such as shown in Fig. 3, Plate IV, are obtained. Because of the great path difference, this instrument is also much used for the analysis of a single spectral line.

92. Michelson's Interferometer. — At this stage the student will scarcely need to be told that an interferometer is an instrument which makes use of the phenomenon of interference to make various measurements. One of the most interesting as well as one of the most useful of such instruments is one due to Michelson.

In the standard form the arrangement is after the manner illustrated in Fig. 127, where A, B, C and D are glass plates, A and B with faces plane and parallel, the faces M_1 and M_2 of plates D and C being optically plane and well silvered. M_1 and M_2, therefore, are excellent optical plane mirrors, while, in addition, the surface M of plate A is frequently half silvered, although this is not necessary.

Suppose now that a represents a single ray from a broad source, which, on striking the surface M is partly reflected as ray 1,

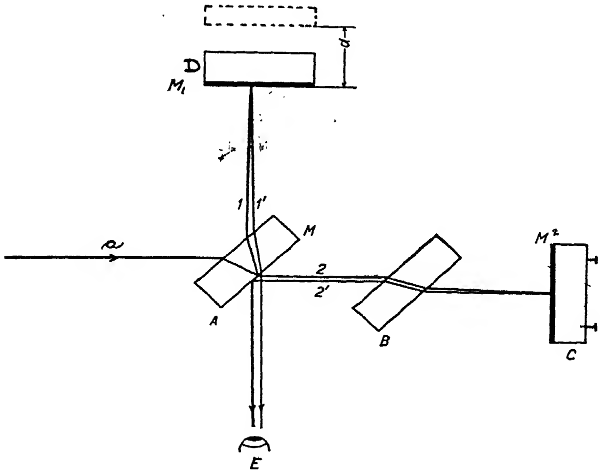


FIG. 127.

partly transmitted as ray 2. On striking mirror M_1, ray 1 is reflected as 1′, subsequently falling on the eye at E, while ray 2 after traversing plate B and after reflection at mirror M_2, retraces its path as ray 2′, finally, after another reflection at M, travelling to the eye along with 1′. Since 1′ and 2′ originally came from the same disturbance, we once more have conditions suitable for interference. When, therefore, a broad source of light is used, in the whole field of view fringes are seen. For the proper adjustment and use of the instrument, the plane of mirror M_2 can be

altered by adjusting screws, while the mirror M_1 can be moved parallel to itself over a considerable distance. The plate B, equal in thickness to A, is inserted between M and M_2 for the purpose of making the optical paths of the two interfering beams similar. It will be noticed that, without it, ray 1 and 1′ after reflection at M would traverse a glass plate twice, ray 2 and 2′ not at all, whereas, by the insertion of plate B, an equivalent thickness of glass is also traversed by ray 2′ and 2.

Suppose, now, that such an instrument has been adjusted to give good fringes with an approximately monochromatic source, a sodium flame for example. If, then, by means of a slow motion screw the mirror M_1 is moved a small distance d cm., then the optical path of ray 1 is increased by $2d$ cm. Since, during the time that the path difference is increased by one wave-length, at any given place a bright fringe changes from bright to dark and back again to bright (a shift of one fringe), it follows that the slow motion of M_1 will cause a continuous shift of the fringes. If the change from bright to bright takes place n times when the mirror is moved d cm. then

$$n\lambda = 2d. \tag{9.06}$$

If, therefore, by means of a micrometer screw the distance d can be measured, we have from this simple relation another means of evaluating λ.

Conversely, and herein we have an important application of interference, if a source emitting a known wave-length is used, the same relation provides us with an exact means of measuring small distances.

An important example of such an application is found in the work of Prof. Michelson on the measurement of the standard metre in terms of light wave-lengths. The standard metre, kept under lock and key at the International Bureau of Weights and Measures near Sèvres, France, and inspected only once in 10 years, defines the unit of length in terms of the distance between two scratches on a bar of platinum — iridium alloy. While there are secondary standards, it is highly important that the length of the metre should be known in terms of an indestructible and perma-

nent unit. Now certain sources emit radiations whose wave-lengths have been shown to be reproducible and constant to a high order of accuracy. One such radiation is the red line of luminous cadmium, whose wave-length in air at 15° C. and 760 mm. pressure is 6438.4696 A. This is about the nearest approach to absolute monochromatism that can be obtained. Using this as a light source, Michelson found the length of the standard metre to be equal to 1553163.5 wave-lengths, to an accuracy of one part in two million.

Fundamentally the method used consisted in the measurement of the distance in terms of a count of fringes. As a count of such a large number would not only be very laborious, but also would increase the possible error of skipping a fringe or two, optical

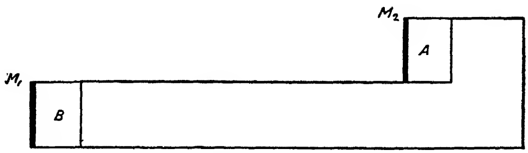


FIG. 128.

sub-units somewhat as illustrated in Fig. 128 were used. In each of these, blocks A and B were placed with silvered faces M_1 and M_2 plane and parallel, while their distances apart were made 10 cm. $\frac{10}{2}, \frac{10}{2^2} \cdot \qquad \frac{10}{2^8}$ cm. or 0.390625 mm. By actual count of fringes the smallest distance was measured. This was then used as a unit of length by means of which (again by the use of the interferometer) the second distance was measured. Proceeding in this way the 10 cm. unit was accurately known in terms of a number of wave-lengths. Finally the length of the standard metre was measured, using the decimetre piece as unit, with the result already given. For full details the student should consult Michelson's "Light Waves and their Uses" or his more recent "Studies in Optics."

Since Michelson carried out this investigation in 1893–1895, in association with Benoît (France), several other similar deter-

minations have been made. Indeed, in 1923 and in 1927 the International Committee on Weights and Measures gave its approval to the principle of using the wave-length of light as the ultimate standard of length, and in 1933 this body appointed a sub-committee to consider further this question. Before the unit of length can finally be defined in this way, careful measurements must be made at the important national laboratories. No report has yet been issued by the U. S. Bureau of Standards but two important accounts of work done at the National Physical Laboratory, in England, have been published. In the second of these, printed in 1934, Sears and Barrell, the experimenters, give the final result, after some ten years work, as 1 metre = 1,553,163.76 wave-lengths of the cadmium red line, in air at 15° C., and 760 mm. pressure, free from carbon dioxide. Some idea of the extreme accuracy necessary in such work is obtained when it is stated that, if the air temperature is 20° C., the result is 1,553,156.33 wave-lengths.

The authors of this paper have collected the results of four important determinations which have been made, (*a*) by Michelson and Benoît in 1893–1895; (*b*) by Benoît, Fabry and Perot (France) in 1905–1906; (*c*) by Watanabe and Imaizumi (Japan), in 1927; (*d*) by Sears and Barrell. If we treat the wave-length of this cadmium line as the unknown and the metre as the known, the mean value of the wave-length in air at 15° C., 760 mm. pressure, containing the normal amount of carbon dioxide, is 6438.4696×10^{-8} cm. It will be recalled (see p. 133) that this is identical with the value used in defining the International Angstrom.

Before leaving the subject of interference a brief summary of its significance may not be amiss. First and foremost, we have been dealing with phenomena which can only be explained in terms of a wave theory of some kind. Moreover, these very phenomena provide us with various methods of measuring the actual wave-lengths of light emitted by a source. Again, once wave-lengths have been evaluated, instruments or optical arrangements by which interference fringes are obtained, enable us to make a great variety of measurements such as the determination

of the index of refraction of a gas, the accurate measurement of small distances, the trueness of surfaces, and the analysis of the structure of a single spectral line. Finally, in the next chapter we shall see how the principle of interference helps us to meet one of the great objections made against the wave theory by the adherents of the corpuscular.

Problems

1. Light from a narrow slit passes through two parallel slits 0.2 cm. apart. The interference bands on a screen 100 cm. away are 0.295 mm. apart. What is the wave-length of the light?

2. The singlet red line of helium has a wave-length of 6678.15 A in air at 15° C. and 760 mm. If the index of refraction of air under the same conditions, for this spectral line, is 1.000276, find its wave-length in vacuum.

3. A beam of monochromatic light is divided; one part is sent through 1 mm. of water, the other part through an air path, so that there may be no relative retardation. What is the air path required?

4. Biprism fringes are produced with homogeneous light of wave-length 6×10^{-5} cm. A thin plate of glass (index of refraction = 1.5) is introduced into the path of one of the interfering beams, upon which the central bright band shifts to the position previously occupied by the fifth. What is the thickness of the glass plate?

5. A narrow filament lamp covered with a red glass, transmitting light for which $\lambda = 6 \times 10^{-5}$ cm., is viewed through a double slit placed immediately before a normal eye. If the distance between the centres of the two slits is 0.3 mm., and if the filament lamp is 3 metres from the slits, describe *quantitatively* the appearance seen.

6. If a thin piece of glass, thickness = 12×10^{-4} cm., is placed in the path of one of the interfering beams in a biprism arrangement, it is found that the central bright band shifts a distance equal to the width of 10 bands. Find the index of refraction of the piece of glass.
Wave-length of light used = 6×10^{-5} cm.

7. A particle is disturbed by six interfering beams, each of the same intensity and every two of which differ in path by $\frac{\lambda}{36}$. Find graphically the ratio of the resultant intensity to that when there is no path difference.

8. At a certain point on a screen, the path difference for the two interfering rays (in an arrangement such as the biprism) is one-sixth of a wave-length. Find graphically the ratio of the intensity at this point to that at the centre of a bright fringe.

9. Why does a vertical soap film, observed by reflected light, become black at the top just before it breaks?

10. Two narrow pieces of plate glass are placed together so as to form a thin wedge. When held in a vertical position and viewed normally by the reflected light from a yellow flame, horizontal interference bands are seen. (*a*) With the eye in any given position, explain by the aid of a diagram the cause of those bands. (*b*) If the distance apart of the bands = 3 mm. and the angle of the wedge = 1×10^{-4} radian, find the wave-length of the light used.

11. A plane vertical soap film is formed over the mouth of a beaker, and illuminated with white light. The light reflected from a narrow portion of it is focussed by a lens on the slit of a spectroscope (with slit horizontal). Describe and explain the appearance observed as the film gets thinner and thinner.

12. Monochromatic light emitted by a broad source of wave-length 6×10^{-5} cm. falls normally on two plates of glass which enclose a thin wedge-shaped air film. The plates touch at one end and are separated at a point 15 cm. from that end by a wire 0.05 mm. in diameter. Find the width between any two bright fringes.

13. Two narrow pieces of plate glass are placed together so as to form a thin air wedge. When held in a vertical position and viewed normally by reflected monochromatic light, horizontal interference fringes 3 mm. apart are observed. If the air space is later filled with water, how far apart will the fringes be?
Index for air — water = 1.33.

14. In a system of Newton's Rings due to a convex lens resting on a plane surface the 25th dark ring is 1 cm. from the centre when sodium light is used. What is the thickness of the air film at that point and what is the radius of curvature of the lens?

15. If the air film in the above problem is replaced by water, index = 1.33, what will be the distance of the 25th ring from the centre?

16. When a Newton's Rings arrangement is used with a source emitting two wave-lengths $\lambda_1 = 6.0 \times 10^{-5}$ cm., $\lambda_2 = 4.5 \times 10^{-5}$ cm., it is found that the n^{th} dark ring for λ_1 coincides with the $(n + 1)^{\text{th}}$ dark ring for λ_2. If the radius of curvature of the curved surface is 90 cm., find the diameter of the n^{th} dark ring for λ_1.

17. If the diameter of the n^{th} dark ring in a Newton's Rings arrangement changes from 1.2 cm. to 1 cm. when the air space is replaced by a transparent liquid, find the index of refraction of the liquid.

18. A film of glass of index of refraction of 1.54 is introduced in the path of one of the interfering beams of a Michelson interferometer and causes a displacement of 20 fringes of sodium light across the field. What is the thickness of the film?

19. A Michelson interferometer is adjusted until good fringes are obtained with monochromatic light. When the movable mirror is shifted 0.015 mm., a shift of 50 fringes is observed. What is the wave-length of the light used?

20. When the movable mirror of a Michelson's interferometer is shifted a certain distance, and a sodium flame is the source of light, 200 fringes are observed to move past a given point in the field of view.

Explain the cause of the fringe movement and calculate how far the mirror was moved.

λ for sodium light = 0.0000589 cm.

21. A source of light emits two wave-lengths $\lambda_1 = 5890$ A, $\lambda_2 = 5896$ A. Interference fringes are observed with a certain arrangement when the paths of the interfering beams are exactly equal. How much will the path difference have to be increased so that a bright fringe for λ_1 coincides with a dark fringe for λ_2?

22. A double slit (with centre to centre = 0.3 mm.) is placed in front of the objective of a telescope which has been previously focussed on two very narrow bright sources of monochromatic light at a distance of 10 metres from the objective. How far apart are the two sources if the bright fringes arising from one fall on the dark fringes arising from the other? (Take mean $\lambda = 6200$ A.)

23. A spectrometer (without prism) is adjusted in the usual way with a narrow source slit illuminated with red light. When a double slit whose openings are 0.5 mm. apart (centre to centre) is placed before the objective of the telescope, interference fringes are observed. If the angular width of 5 fringes is found by measurement to be 25 min. find the wave-length of the red light. Explain your solution with the aid of a good diagram or diagrams illustrating the complete optical arrangement.

24. A thin wedge-shaped air film is formed between two optically plane glass surfaces which touch at one end and are separated at the other, at a distance of 10 cm., by a flake of mica 0.04 mm. thick. The film is viewed normally by reflected light of wave-length 6.0×10^{-5} cm. (i) How much must the plates be separated at the place where they originally touched so that a shift of 6 fringes is observed at this place? (ii) What is the distance between any two successive bright (or dark) fringes observed on the face of the film in the new position of the plates?

25. Two interfering rays of equal intensity unite to form a resultant whose intensity is 50% of the intensity obtained when there is no phase difference between the two rays. Find the phase difference between the two rays.

26. A spherical glass surface resting on a plane glass surface forms a typical Newton's Rings arrangement (centre black). Prove that the

difference of the squares of any two successive bright rings is a constant.

27. In a Newton's Rings arrangement, find, for $\lambda = 6.40 \times 10^{-4}$ mm. the *resultant* phase difference between the two interfering rays, at a place where the thickness of the air film is 0.00656 mm.

28. Monochromatic light of wave-length (in air) $= 6.24 \times 10^{-5}$ cm. is reflected approximately normally from the two plane faces enclosing a thin film of water. If the index of water for this wave-length is 1.34, find (in degrees) the phase difference between the two interfering rays at a point on the face of the film where the thickness is 2.8×10^{-3} mm.

29. Prove that the phase difference between the resultants of each of the apertures of a double slit is equal to the phase difference between rays from the centres of each of the apertures.

30. A double slit whose width (centre to centre of slits) is 0.5 mm. is placed in front of a telescope and the resulting fringes (when a small distant bright source is viewed) are observed in the focal plane of the objective, whose focal length is 50 cm. If the mean value of λ is 5.5×10^{-5} cm. find the ratio of the intensity at a point at the centre of a bright fringe to that at a point 0.10 mm. away from the centre (both points in the focal plane).

31. Given a double slit, with known distance b between the centres of its openings, and a spectrometer, describe with diagram and explanation how you could measure the mean wave-length emitted by a sodium flame.

32. Any device which may be used for obtaining interference fringes may also be used as a means of measuring the wave-length of the light used. Discuss this statement.

33. Explain what is meant by a *shift of interference fringes*, with reference to either a Rayleigh or a Michelson interferometer. Give one example of an application of such a shift, outlining the necessary calculations.

34. Make a diagram showing how Fresnel mirrors enable one to obtain two sources suitable for realizing interference fringes. Show how the width of the fringes varies with the angle of inclination of the faces of the two mirrors.

35. Explain why fringes of equal thickness cannot be observed by reflection from the two surfaces of a piece of ordinary thick plate glass.

36. Two optical arrangements are used to obtain interference fringes with monochromatic light. In the first there are two interfering beams of equal amplitude. In the second there are eight beams and the arrange-

ment is such that the phase difference between each successive pair is the same.

Find in each case, the ratio of the intensity at the centre of a bright fringe to the intensity at a point to one side of the centre, when the point is so chosen that the phase difference between the two interfering beams, or between each successive pair, is $10°$. (Use a ruler and a protractor.)

CHAPTER X

DIFFRACTION

93. In Chapter I it was pointed out in some detail that the *apparent* impossibility of explaining rectilinear propagation on a wave theory was so serious an obstacle to its acceptance that Newton and many other scientists gave their support to the corpuscular theory. In the chapter preceding this one we have given convincing evidence that a wave theory of some kind is necessary for the interpretation of interference phenomena, but as yet no explanation has been given of the fact that light for all practical purposes travels in straight lines. In the present chapter we shall endeavour to show, (1) that, strictly speaking, light does *not* travel in straight lines, (2) that by combining a modification of Huygens' Principle with the principle of interference, after the manner of Fresnel, approximate rectilinear propagation results *for a wave-motion in which the wave-lengths are very small in comparison with the dimensions of ordinary obstacles and openings.*

Phenomena which make it evident that light, in certain circumstances, is not travelling in straight lines are said to be examples of diffraction, a name we owe to Grimaldi (Italy, 1618–1663), its discoverer. When an opaque obstacle such as AB, Fig. 129, is placed in the path of a beam of light emerging from a small hole S, Grimaldi showed not only that the width of the shadow cast on a screen is larger than the width CD it would have if light travelled in straight lines, but also that this wider shadow is bordered by coloured bands. The pioneer work of Grimaldi was repeated and extended by Newton, who tried to explain the phenomenon in terms of the repulsion of light corpuscles by the obstacle, without realizing that it was something to be expected on a wave theory.

There are many other ways of showing that light does not travel in straight lines. If, for example, an adjustable slit P, Fig. 130, is placed in the path of a beam emerging from a narrow

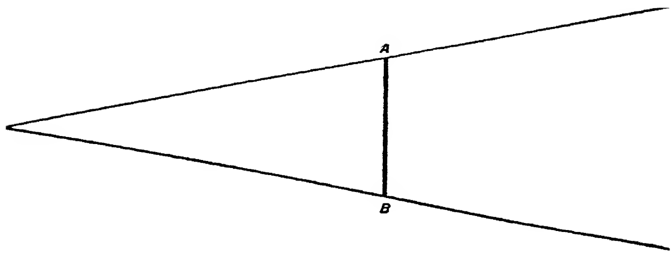


FIG. 129.

source slit S, and the light pattern on a screen is observed, it is found, after a certain width of P is reached, that the *narrower* P is made, the *wider* the band of light MN becomes. For a very narrow slit, although the emergent light is faint, it covers an

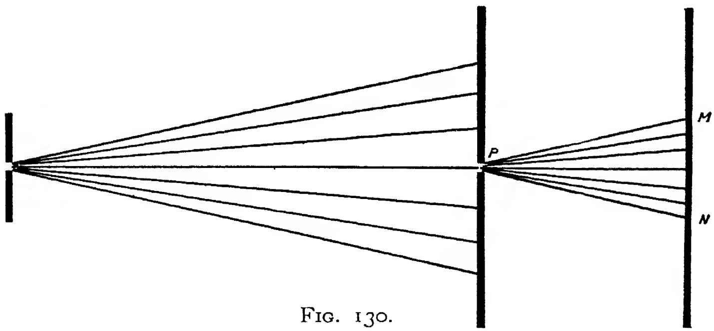


FIG. 130.

area hundreds of times greater than that which would be illuminated if light travelled in straight lines from the source to the screen.

Again, if a small perfectly circular object D, Fig. 131, is placed in the path of the beam from S, instead of having a uniformly dark shadow MN, there is a *bright* spot at the centre of the cir-

cular shadow. An actual photograph of the appearance will be found in Fig. 2, Plate V.

Later we shall have occasion to discuss other illustrations of the point we are trying to emphasize, but these should be sufficient to make it clear that it is not difficult to show by simple

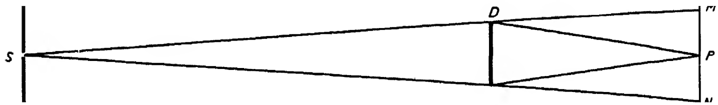


FIG. 131.

arrangements that light does not always even apparently or approximately travel in straight lines. It is natural then to ask, What are the general conditions necessary for making diffraction phenomena evident, and, once again, why is it that, generally speaking, light does appear to travel in straight lines?

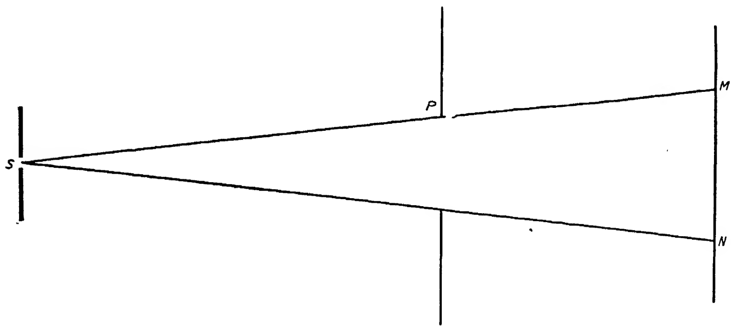


FIG. 132.

94. A suggestive fact which will help in the answer to these questions is found when the slit P, Fig. 130, is made very wide, as represented in Fig. 132. It is then observed that the light patch on the screen has just about the width to be ex-

pected from the laws of geometrical optics (which postulate rectilinear propagation), although here again a close examination shows that at the edges M and N one or two narrow light and dark bands can be seen. Experiment then suggests that the larger an opening (or obstacle), the more nearly may we expect rectilinear propagation. Let us see what can be deduced about the matter from theoretical considerations.

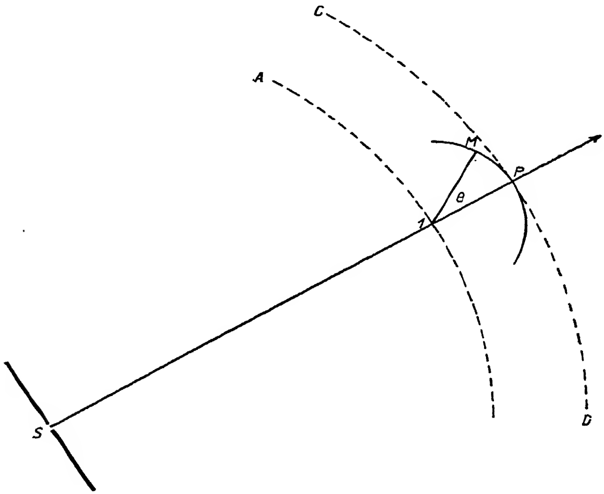


FIG. 133.

Suppose we postulate again, as on p. 48, Chapter III, that every particle in a wave-front is the source of a secondary wavelet, which, however, we shall now consider to be effective over the whole of its wave-front, and not simply at the small area tangent to the enveloping surface.

If then the small circle in Fig. 133 represents a wavelet starting from particle 1 on a wave-front which originally came from the source S, we are assuming that the wavelet is effective not only at P, where CPD is the enveloping surface representing a later position of the wave-front, but also at other points such as M. We shall postulate further that the *amplitude* of the disturbance

at points on the wavelet is less the greater the angle $1M$ makes with the ray $1P$, in other words, that it varies with the obliquity.

Bearing these things in mind, we see that if well defined wave-fronts from a small source S fall on an opening P, Fig. 134, the intensity at any point Q on a screen receiving the emergent light will be the resultant of all the elementary disturbances coming to it from all the sources in the plane of the opening sending out wavelets. The problem, then, becomes one of superimposing a

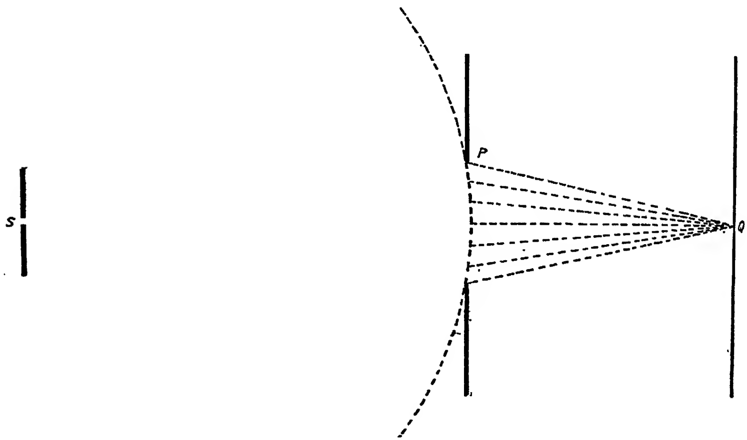


FIG. 134.

large number of disturbances, and therefore is one to which we can apply the principle of interference. The solution is not always easy, and fairly extended mathematical analysis is needed for treating the problem generally with any attempt at exactness. Much, however, can be learned by confining our attention to simple cases and by making use of graphical methods.

95. Half-Period Elements. — First of all, then, let us consider that the point source S is sufficiently far away that we are concerned with plane wave-fronts of monochromatic light. Then let us find the resultant amplitude at a point P, Fig. 135, due to

an aperture limiting the exposed part of the wave-front to AB. As a first step in the solution it is convenient to subdivide the wave-front into a number of elemental areas. To do so, drop a perpendicular $PM_0 = b$, from P to the wave-front; then with M_0 as centre draw on the wave-front a circle of radius M_0M_1,

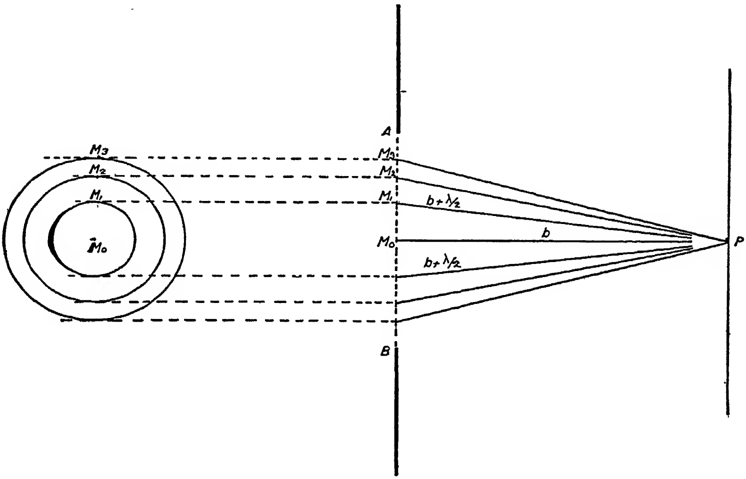


FIG. 135.

the magnitude of M_0M_1 being such that the distance $PM_1 = b + \frac{\lambda}{2}$. Next, describe a second circle of radius M_0M_2, such that $PM_2 = b + 2 \cdot \frac{\lambda}{2}$, and so on. In this way the whole of the wave-front is subdivided into areas called *half-period elements or zones*, of such a size that the distance from the outer edge of any one to the point P is greater than that from the outer edge of the preceding one by half a wave-length.

The areas of these half-period elements are equal *when the wave-length in question is small in comparison with the dimensions of ordinary openings and obstacles.*

$$
\begin{aligned}
\text{Area of first elemental area} &= \pi M_0M_1^2 \\
&= \pi(PM_1^2 - PM_0^2) \\
&= \pi\left\{\left(b + \frac{\lambda}{2}\right)^2 - b^2\right\} \\
&= \pi b\lambda + \pi\lambda^2 \\
&= \pi b\lambda,
\end{aligned}
$$

if λ is small compared with b.

Again, the area of the first and the second elements

$$
\begin{aligned}
&= \pi M_0M_2^2 \\
&= \pi(PM_2^2 - PM_0^2) \\
&= \pi\{(b + \lambda)^2 - b^2\} \\
&= 2\pi b\lambda.
\end{aligned}
$$

$\therefore$ area of the second element $= \pi b\lambda$.

Proceeding in this way we see generally that, if λ is sufficiently small compared with b that we may with little error neglect λ^2, then the area of each half-period element $= \pi b\lambda$. For $\lambda = 6 \times 10^{-5}$ cm. and $b = 50$ cm., it is easy to show that the error in using this expression for the area is 0.00003 per cent. for the first element and 0.006 per cent. for the 100th.

Actual Size of Half-Period Elements. — Suppose we are dealing with light for which $\lambda = 5 \times 10^{-5}$ cm., and that the screen is placed at a distance of 50 cm. from the aperture, then the area of each element

$$
\begin{aligned}
&= \pi b\lambda \\
&= 3.14 \times 50 \times 5 \times 10^{-5} \text{ sq. cm.} \\
&= 0.00785 \text{ sq. cm.}
\end{aligned}
$$

As ordinary areas go, therefore, half-period elements are extremely small. A circular aperture whose radius is 1 cm. contains 400 in this particular case.

The student must not fail to realize, however, that the actual area of these hypothetical elements depends both on the wavelength and on the distance of the screen or eyepiece from the wave-front.

Although the area of each element may be taken as $\pi b\lambda$ to a high order of accuracy, in considering the various factors which affect the resultant amplitude at P, in Fig. 135, it is not desirable to neglect the very slight increase in the areas of the elements as we go out from M_0. Let us examine this question carefully. The amplitude at P *due to a single half-period element* depends on: (1) the area of the element; (2) its mean distance from P; and (3) the obliquity, that is, the angle between the line joining the element to P and a perpendicular to the plane waves incident on AB, in Fig. 135.

(1) As far as area is concerned, it is reasonable to assume that the amplitude at P is directly proportional to the area. The greater the area, the greater the number of particles sending out wavelets in accordance with the modification of Huygens' Principle we are discussing.

(2) As far as distance is concerned, we know from the work discussed in section 21, Chapter II, that the amplitude falls off inversely as the distance.

If we consider the *combined effect* of both area and distance, it is easy to show that the contribution of each element is *exactly the same*. Here is the proof.

From the above we can state that the amplitude at P due to a single element is proportional to $\dfrac{\text{area}}{\text{distance}}$.

But the area of the nth element

$$= \pi\left\{\left(b + n\frac{\lambda}{2}\right)^2 - b^2 - \left(b + \overline{n-1}\,\frac{\lambda}{2}\right)^2 + b^2\right\}$$

$$= \pi\lambda\left\{b + \left(2n - 1\right)\frac{\lambda}{4}\right\}.$$

Also, the mean distance of the nth element from P

$$\frac{1}{2}\left(b + n\frac{\lambda}{2} + b + \overline{n-1}\,\frac{\lambda}{2}\right)$$

$$= b + \left(2n - 1\right)\frac{\lambda}{4}.$$

Hence,

$$\frac{\text{area}}{\text{distance}} \quad \frac{\pi\lambda\left\{b + \left(2n - 1\right)\dfrac{\lambda}{4}\right\}}{b + \left(2n - 1\right)\dfrac{\lambda}{4}}$$

$$= \pi\lambda, \text{ a constant for all values of } n.$$

(3) There is left only the obliquity factor. As we have already postulated, the amplitude at points in a wavelet originating in a secondary source decreases with increasing obliquity. We conclude, therefore, that, as we go outwards from M_0 in Fig. 135, the amplitude due to each successive element decreases steadily, although extremely slowly because of the scarcely detectable difference in the obliquities of two neighboring elements.

In considering the *resultant* amplitude due to all elements making their contribution at P, differences in phase are of primary importance. Because of the way in which the half-period elements are constructed, the phase difference between any two successive zones is π radians or 180°. Our original problem, therefore, may now be re-stated as follows:

Find the resultant amplitude for n S.H.M., all in the same straight line, whose amplitudes decrease extremely slowly and between every successive two of which there is a phase difference of π.

Let d_1 = amplitude of disturbance at P due to the first element alone,

d_2 = amplitude of disturbance at P due to the second element alone,

or, in general,

d_n = the amplitude due to the nth element.

As we have seen, d_1 is *slightly* greater than d_2, d_2 than d_3, d_n than d_{n+1}. Moreover, the difference between any two successive d's becomes less and less the higher the numbers of the elements because of their increasing narrowness as we go out from the first.

In solving this problem we apply the proposition of section 15, Chapter II, obtaining the result that all the amplitudes

(because of the phase difference of π) lie on the same straight line, as in a, Fig. 136. For the purpose of clearness we have separated the individual amplitudes as in b, Fig. 136 and c, Fig. 136, but it must not be forgotten that they all lie along the same straight line. The resultant amplitude is represented by BA, and it will be noted that b, Fig. 136 applies to odd values of n, c, Fig. 136 to even.

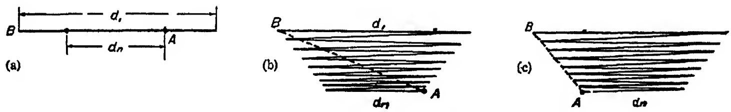


FIG. 136.

The actual value of BA, the resultant, which we shall henceforth represent by D, will be found in a few special cases.

Case I. **Circular Aperture Containing Only a Few Elements.** — Suppose the aperture AB, Fig. 135, is circular and is so small that $n = 1$ (the student should find its actual radius for particular values of b and λ). In that case, $D = d_1$.

When $n = 2$, $D = d_1 - d_2$, or is practically zero. In other words, the intensity of the light at P when the opening contains two half-period elements is a minimum, or P is dark.

When $n = 3$, $D = d_1 - d_2 + d_3$, that is, is very nearly equal to d_1 or d_3. In this case, the intensity at P is a maximum, or P is bright.

We see then generally, that if a small circular opening contains a few elements the number of which is odd, the intensity at P is a maximum, whereas if the number is even, the intensity is a minimum. The truth of this can readily be verified by direct experiment, if a circular opening, of diameter 1 or 2 mm. is illuminated with light coming from a distant pin-hole source and the emergent light is received on a screen or viewed in an eyepiece. By gradually altering the distance b from the eyepiece to the aperture, it will be found that the centre of the light pattern is alternately bright and dark. Since the area of a half-period element depends on b, as b is altered the opening of fixed area

must contain alternately an odd and an even number of elements. Actual photographs of the appearance are reproduced in Fig. 1, Plate V, where Fig. *b* shows a bright centre, Fig. *a*, a dark one.

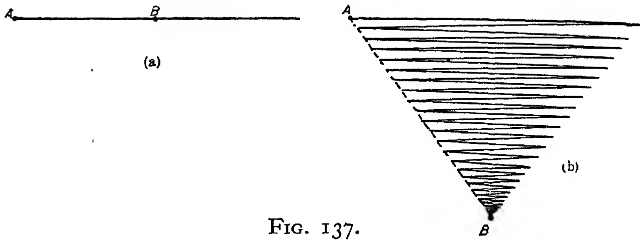


FIG. 137.

Case II. **Large Aperture Containing Very Many Zones.** — If the number of exposed elements is sufficiently large, then ultimately the value of d_n will become negligibly small, and the figures corresponding to *a*, and *b* or *c*, Fig. 136 will be as shown in *a* and *b*, Fig. 137. This figure shows at once that

$$D = \frac{d_1}{2}. \tag{10.01}$$

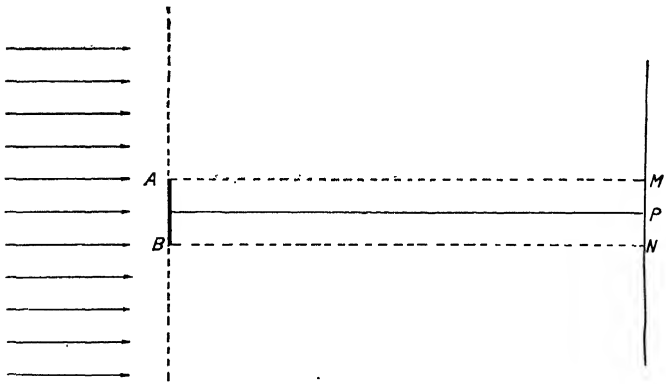


FIG. 138.

This important result tells us that when the exposed portion of the wave-front contains a very large number of half-period elements, or, in other words, when the dimensions of an opening

are large compared with the wave-length, the resultant intensity at a point such as P is the same *as if* the light came directly from an opening smaller than the first half-period element.

Case III. **Circular Obstacle.** — Consider the intensity at a point P at the centre of the geometrical shadow MN, Fig. 138, cast by a circular object AB illuminated by plane waves of monochromatic light. According to the view we are presenting, disturbances from all the particles in the exposed portion of the wave-front arrive at P. What is the resultant in this case?

Let the object cover k elements. Then, since the $(k + 1)$ element is the first from which P receives light, the problem is exactly the same as the one we have considered for a large opening, except that d_{k+1} replaces d_1. The figure for finding the resultant otherwise will be exactly the same as Fig. 137. It follows, therefore, that the resultant amplitude at P for this case is

$$\frac{d_{k+1}}{2}.$$

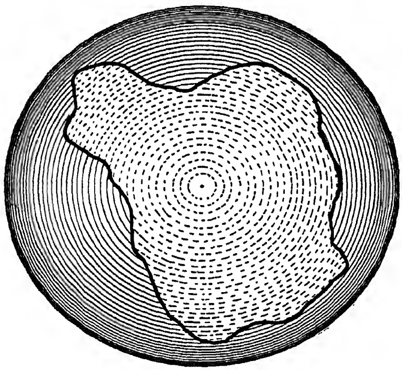


FIG. 139.

If now the object is so small that only a few elements are covered, this resultant amplitude cannot be neglected. This means that there is then a bright spot at the centre of the shadow MN. That this is actually the case has already been pointed out when reference was made to Fig. 2, Plate V. If the object is small and not circular, no such spot will be seen because, in the case of an irregular object, irregular portions of a number of elements will be "exposed" in a random fashion, as Fig. 139 suggests, and no well-defined maximum at P is possible.

Case IV. **Large Obstacle.** — If an object is large, that is, if it is one which covers a large number of elements and therefore

has dimensions large compared with the wave-length, then k is so great that the value of $\frac{d_{k+1}}{2}$ may be neglected. This means that for the majority of ordinary objects (whose dimensions fulfil this condition) points such as P are completely screened. The reason again is *not* because light travels in straight lines from the distant point source, but because the application of the principle of interference shows that the resultant of all the disturbances from the various parts of the exposed wave-front is practically zero. Conditions, however, are the same *as if* light travelled in straight lines from the distant point source and the "direct" rays were cut off by the obstacle.

Approximate rectilinear propagation, then, results from the fact that ordinary openings and ordinary objects have dimensions many times the mean wave-length of light. The screening action of obstacles depends on the wave-length of the disturbance, consequently in the case of sound, where we have to do with wave-lengths of a foot or two, it is not surprising that we do not observe shadows. To obtain shadow effects, objects of great dimensions are necessary.

96. Diffraction effects prove that light does "bend around corners," and in general behaves in this respect like any other kind of wave motion, the difference being in degree, not in kind. But even diffraction effects are not observed unless the object or aperture is properly illuminated. In the examples we have given, the distant source was always so narrow that we were able to consider well-defined wave-fronts in the plane of the aperture or opening. If, however, a broad source is used, or if we are dealing with the general illumination due to diffuse light, no such effects are observed even in the case of small objects. The reason is much the same as has been given to explain why double slit fringes cannot be obtained with a broad source. Every elemental source into which the wide one may be subdivided will give rise to its own diffraction pattern. The superposition of all of these makes the formation of maxima or minima impossible.

97. Zone Plate. — An interesting example of diffraction, as well as a confirmation of the above theory, is found in the zone plate. This is a simple device whose construction is based on the fact that, since the areas of the half-period elements are equal, their radii are in the ratios $1 : \sqrt{2} : \sqrt{3} : \sqrt{4}$: etc.

Suppose, then, that a set of circles, whose radii increase according to this law, are drawn on a sheet of white paper, and that every second area is blackened. If this pattern is then photographed, the resulting plate will give a system of areas corresponding to the half-period elements, every second one of which is opaque. If this zone plate, as it is now called, is placed in the path of plane waves of monochromatic light and the emergent light allowed to fall on a screen, there will be a certain position of the latter (that is, a certain value of b) for which the hypothetical half-period elements coincide with the actual areas of the zone plate. In that case, when the even numbers are opaque,

$$D = d_1 + d_2 + d_3 + \cdot \cdot \cdot,$$

since the phase difference between every successive exposed area is 2π. Again the graphical solution gives a straight line but this time the lines representing individual amplitudes all point in the same direction, lying end to end. A maximum of such great intensity results that the zone plate is sometimes spoken of as having a focal length equal to this particular value of b.

98. More General Treatment. — Hitherto we have dealt with only the resultant amplitude due to each half-period element. Actually, as we go from particles just on the inner edge of a single element to the outer edge of the same element, the phase difference must gradually change. It is our purpose now to take into consideration this gradual change of difference in phase between disturbances from particles in a single element. To do so, consider the first element and imagine it subdivided into a large number of very small elemental ring areas subject to the condition that the distance to the point P, Figs. 135 and 140, from the edge of any ring differs from the distance of the corresponding edge

of the next by the same amount. There will then be the same phase difference between the disturbances from any two succes-

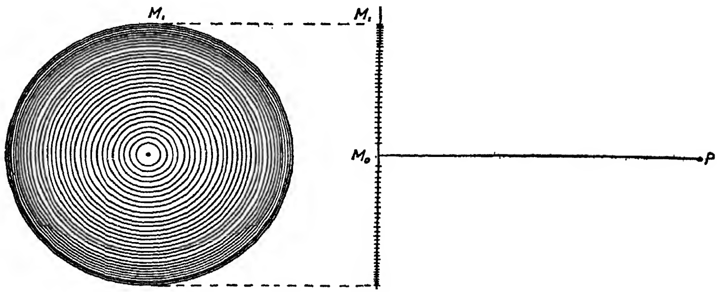


FIG. 140.

sive elemental areas. Moreover, it may readily be proved, just as we did for the half-period zones themselves, that all these sub-areas are equal. Therefore, if a is the amplitude at P due

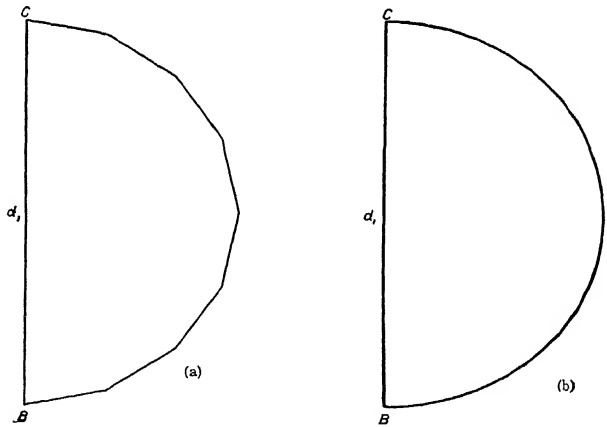


FIG. 141.

to the first of these elemental areas about M_0, the amplitude of the last (in the neighbourhood of M_1) will differ from a by a negligibly small amount.

Suppose, now, we have n of these sub-areas in a single half-period element such as the first, and we wish to find the resultant disturbance. Since the phase difference between any successive two is the same, since the phase difference between the first and the last is π, and since the individual amplitudes are practically equal, the polygon from which the resultant is found a, Fig 141, will approach a half circle more and more closely, as the number of sub-elements (which may be as large as we please) is increased. In the limit, then, b, Fig. 141 is the graphical solution from which we see that the resultant of the first half-period element $= BC$,

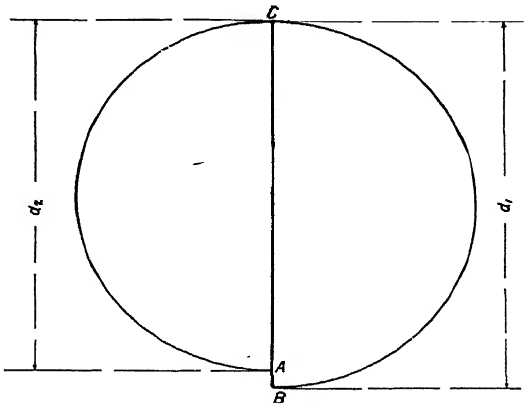


FIG. 142.

that is, d_1 in the above notation. Moreover, although this is a result of which use will not be made, $d_1 = \dfrac{2na}{\pi}$, since na is the length of half the circumference of a circle of diameter d_1.

If we proceed in similar fashion to find the resultant of the first two zones, then a second half-circle will be added to b, Fig. 141, whose diameter $CA = d_2$, Fig. 142, practically could not be distinguished from BC. Theoretically, since there is a very slight falling off in amplitude even in going out from the centre a distance equal to two zones, the half-circles are not perfect and the point A does not quite coincide with B. Once more, however, the re-

sultant of the two zones, equal to BA, that is, $d_1 - d_2$, is negligibly small.

When a large number of elements is considered it is not difficult to see that the graphical solution by this method is given by Fig. 143. The final resultant is then AB, which is equal to $\frac{BC}{2}$ or $\frac{d_1}{2}$, just as in the other method.

99. Other Examples of Diffraction. — Some very interesting, as well as beautiful, examples of diffraction are obtained when a pin-hole distant source is replaced by a very narrow slit, and nar-

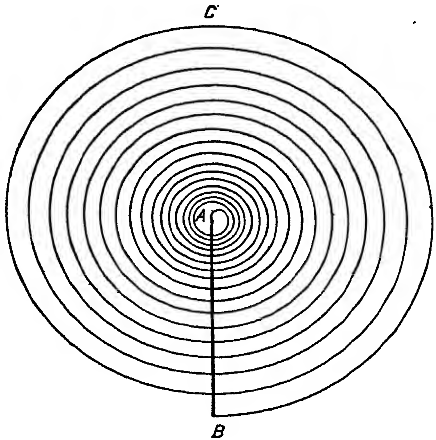


FIG. 143.

row rectangular apertures, or objects such as wires and needles are used. For example, Fig. 5, Plate V, is a photograph of the diffraction pattern of a narrow aperture, with width and screen distance such that the centre is dark. On the other hand Fig. 3, Plate V, shows the pattern obtained with a narrow wire, conditions being such that there is a bright line down the centre of the shadow.

In considering diffraction problems where we have to do with a narrow source slit, we may adopt a graphical method very

similar to the one we have used for a pin-hole source. In the former case, however, if the slit is sufficiently long, we must consider cylindrical wave-fronts rather than spherical. When, therefore, the resultant intensity is wanted at points along a line running through the point P, Fig. 144, and parallel to the slit, we sub-divide the wave-front into strips rather than zones. To do so the same general method is followed as for half-period zones. Thus, $PM_0 = b$ is drawn perpendicular to the wave-front; a line through M_1 defining the outer edge of the first strip in the upper

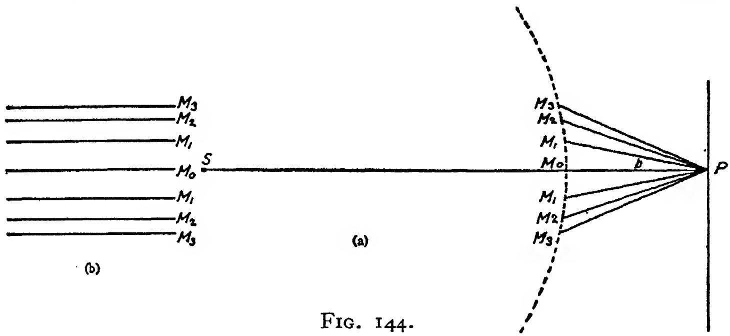


FIG. 144.

half of the wave-front is drawn subject to the condition that $PM_1 = b + \frac{\lambda}{2}$; the upper edge of the second strip is defined by a line through M_2 such that $PM_2 = b + \frac{2\lambda}{2}$, and so on for the others. The lower half is subdivided similarly so that looking at the wave-front "face on," the very exaggerated appearance is somewhat as shown in b, Fig. 144.

Now while these strips are extremely narrow, their areas, unlike those of the half-period zones, are not equal, but decrease, at first fairly rapidly, as the distance from M_0 increases. The result of this is that, when we find the resultant at P by a graphical method similar to that exemplified in Fig. 143, we obtain for the solution a diagram like that given in Fig. 145. While mathematical analysis is needed to determine the exact nature of this figure, called *Cornu's Spiral*, its general features should be clear

enough by analogy with the solution for a pin-hole source. Thus, since the difference in phase between disturbances from the corresponding edges of two successive strips is π, OM represents d_1, the amplitude contributed by the first strip in the upper half, OM' the corresponding d_1' for the first strip of the lower half; ON, d_2 for the first and second strips of the upper, ON', d_2' for the first and second of the lower, and so on.

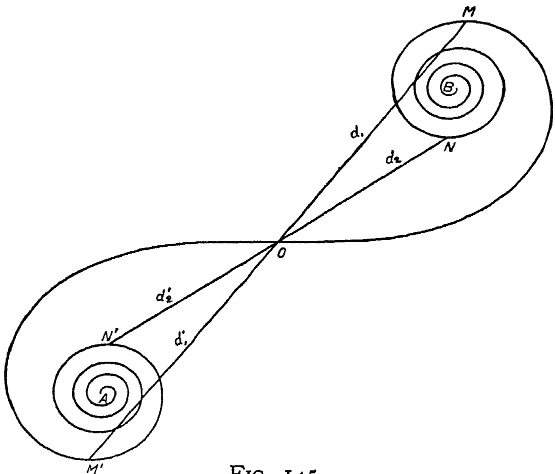


FIG. 145.

If, then, the whole wave-front is exposed the resultant amplitude is represented as always (see again section 15, Chapter II) by the line joining the beginning and end of the figure formed by the individual amplitudes — in this case, by AB. If, however, a narrow aperture is used covering only the first strip in each half, then the resultant amplitude is $MM'(2d_1)$. If such an aperture is widened so that it now includes two strips in each half, then the resultant is equal to $NN'(2d_2)$. Thus, for this second width, the intensity at P, the centre of the diffraction pattern, is considerably less than for the first one.

An interesting example of this type of diffraction has to do with a straight edge. Suppose S, Fig. 146, represents a long

narrow source slit running perpendicular to the plane of the paper, AO an obstacle whose straight edge A is parallel to the

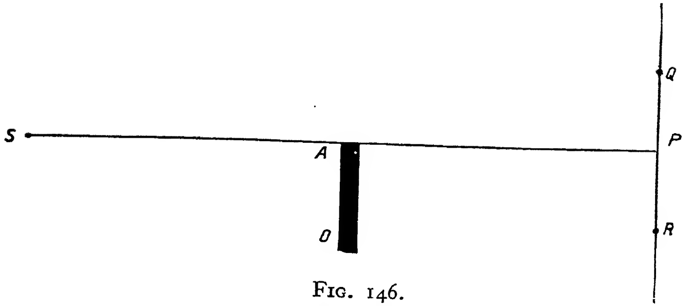


FIG. 146.

slit, and QPR a screen, P being the line which separates light from darkness according to the laws of geometrical optics (that is, exact rectilinear propagation). Consider now the intensity,

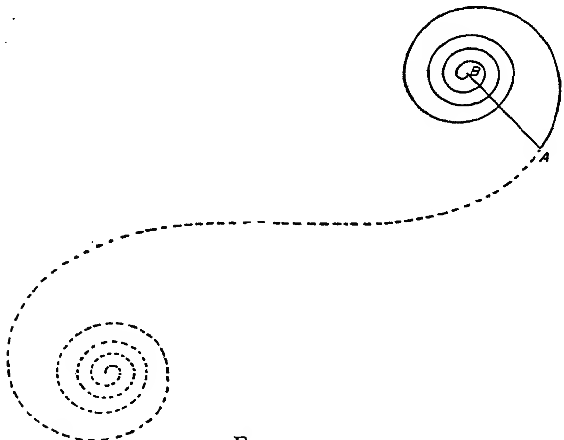


FIG. 147.

(*a*) at the point P (in reality at points along a line running through P and parallel to the slit); (*b*) at points such as R in the shadow region below P; (*c*) at points such as Q in the light region above P.

Since at P the lower half of the spiral is covered by the obstacle, exactly the upper half of the wave-front being "exposed," the resultant amplitude at P is represented by OB, Fig. 145.

For all points such as R, it will be seen that not only the whole of the lower half of the spiral is covered, but that smaller and smaller portions of the upper half are exposed, the farther R is taken from P. The resultant amplitude, therefore, which in Fig. 147 is represented by BA, is steadily decreasing as points are taken farther and farther into the shadow region.

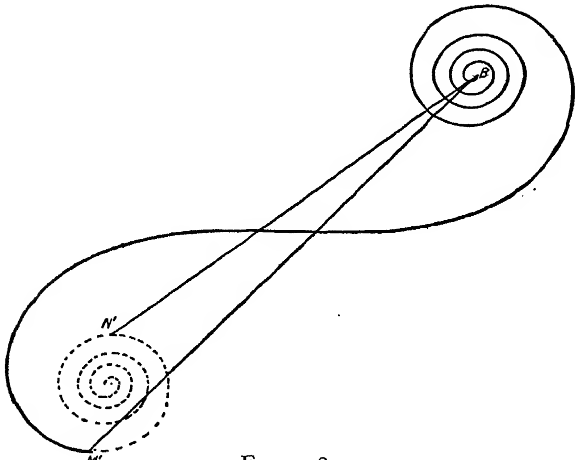


FIG. 148

Turning next to points such as Q, Fig. 146, above P, we note first of all that all of the upper half of the wave-front is exposed, together with a portion of the lower whose magnitude increases the farther Q is taken from P. Thus, when Q is in such a position that the first strip of the lower is added to the upper half, the resultant is represented by BM', Fig. 148, while when Q is sufficiently far removed from P that the first two strips of the lower are added, the resultant is BN' in the same figure. We see, then, that as points are taken farther and farther away from P in the light region, we have alternately maximum and minimum values

PLATE V

(a) (b)

Fig. 1. Diffraction pattern for circular aperture. (*a*) centre dark, (*b*) centre bright.

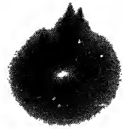


Fig. 2. Shadow of circular object, showing bright centre.

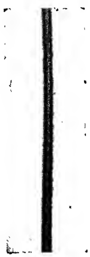


Fig. 3. Diffraction pattern for narrow obstacle, showing centre bright.

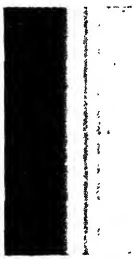


Fig. 4. Diffraction pattern for straight edge.

Fig. 5. Fresnel diffraction pattern for rectangular slit, showing centre dark.

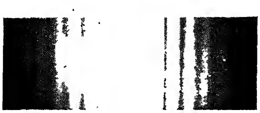


Fig. 6. Fraunhofer diffraction pattern for narrow rectangular aperture.

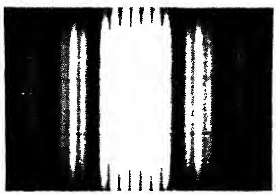


Fig. 7. Fraunhofer diffraction pattern for double slit, showing single slit diffraction fringes with double slit interference fringes.

for the resultant amplitude. A careful use of Cornu's Spiral in this way enables one readily to plot an intensity curve of the general nature shown in Fig. 149, where AB corresponds to the square of AB of Fig. 147; OB to the square of OB of Fig. 145; and BM', BN' to the squares of BM', BN' of Fig. 148. The actual appearance of such a diffraction pattern is reproduced in Fig. 4, Plate V.

The following problem is given as a further illustration of the use of the spiral.

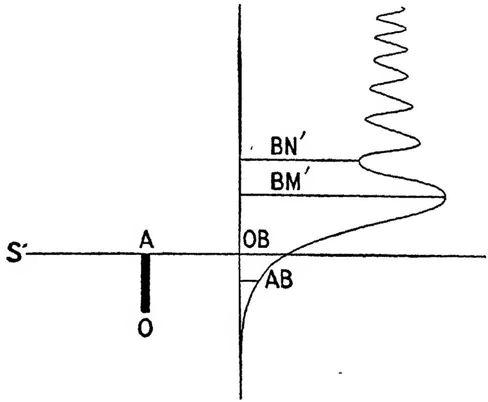


FIG. 149.

When using an arrangement like that of Fig. 146, prove that the intensity at P, if the obstacle AO is removed, is four times the intensity at this point with the obstacle in the position shown.

With the obstacle removed, the whole wave-front is effective, and the resultant amplitude at $P = AB$, in Fig. 145.

As already noted, with the obstacle in position, the resultant amplitude at $P = OB$, in Fig. 145.

Hence, the required ratio

$$= \frac{AB^2}{OB^2} = \frac{4OB^2}{OB^2} = \frac{4}{1}.$$

100. Fresnel and Fraunhofer Diffraction. — There are two general classes into which diffraction phenomena may be divided. In the first, to which the name Fresnel is applied, the intensity at any point is the resultant of disturbances coming directly to that point from all parts of the exposed wave-front. This is the kind we have been considering in this chapter. In the second, designated by the name Fraunhofer, a lens is placed beyond the aperture or obstacle, and the diffraction pattern is examined in the plane where a sharp image of the source would be formed in the absence of the aperture or obstacle. Since the source is generally a distant one, this plane is never far from the focal plane of the lens, and in all cases which we shall consider in the next chapter, is identical with that plane.

CHAPTER XI

FRAUNHOFER DIFFRACTION

101. Three important cases of this type of diffraction will be discussed, (1) a single opening, (2) two openings (the double slit), (3) many openings (the grating). In all cases we shall assume that plane waves fall on the openings, and hence that

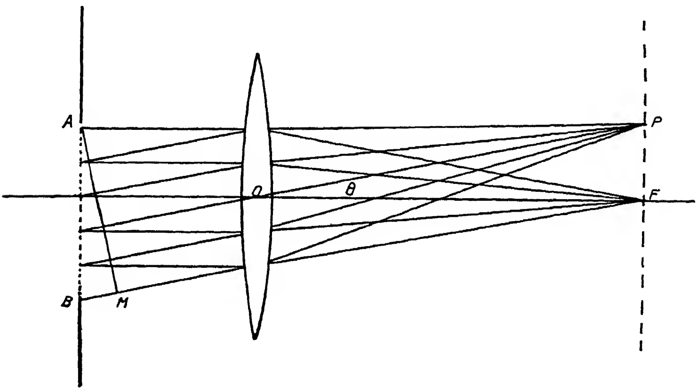


FIG. 150.

the diffraction pattern is observed in the focal plane of a lens. It follows that, for visual observation, a spectrometer (adjusted in the usual way described in Chapter VI, section 58) provides an excellent arrangement if the aperture is placed in the parallel beam leaving the collimator.

Single Rectangular Aperture. — Suppose, then, that a narrow rectangular aperture is placed in the path of plane waves before the objective of a telescope. Applying the same ideas as in the last chapter, we see that there will be a resultant disturbance at points such as P, Fig. 150, in the focal plane of the objective. There will be this difference, however, that for each point such as P we have to deal with a large number of *parallel* rays incident

on the lens, and not, as in Fig. 134, for example, with rays going directly to Q from the elemental sources in the plane of the opening. For this reason, two important conclusions at once follow. (1) Since the path length for all rays parallel to the axis of the lens and focussed at F is the same, the centre of the diffraction pattern is always a maximum, and not sometimes bright, sometimes dark, as in Fresnel phenomena. (2) Since, as already noted, for every point such as P the rays incident on the lens

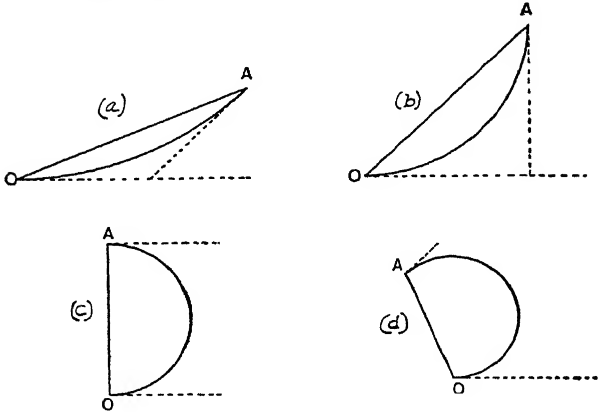


Fig. 151.

have all the same obliquity, the amplitude contributed by each elemental disturbance is exactly the same.

To find the resultant amplitude at P, we can subdivide the aperture into as many elemental sources as we like, and apply the same graphical method as already has been used so frequently. In the case we are considering, since the amplitude contributed by each elemental source is the same, and since there is always the same phase difference between disturbances from any two successive sources, the figure from which the resultant amplitude is obtained is always an arc of a circle. Thus, if the path difference between the disturbances coming to P from the two edges of the aperture, that is, if BM, Fig. 150, is $\frac{\lambda}{8}$, the resultant amplitude

at P is represented by OA, a, Fig. 151; if $BM = \frac{\lambda}{4}$, by OA, b, Fig. 151; if $BM = \frac{\lambda}{2}$, by OA, c, Fig. 151; and if $BM = \frac{3\lambda}{4}$ by OA, d, Fig. 151. It will be noted that as we go out from F, where, again, the intensity is always a maximum, the resultant must at first steadily decrease, becoming zero when the arc closes to form an exact circle. For this result to be obtained, the phase difference between the disturbance from one edge must differ from that coming from the other by 360°, or the corresponding path difference BM must be exactly one wave-length.

If then θ_1 is the angular distance from F to P when the intensity has dropped to zero, since in Fig. 150 $\angle POF = \angle BAM$, we have the important result

$$\sin \theta_1 = \frac{\lambda}{AB} = \frac{\lambda}{a}, \tag{11.01}$$

where a = width of the narrow aperture.

As we take points still farther away from P, since the phase difference between successive elemental disturbances becomes steadily greater, it is easy to see that the arc forms more than a whole circle, the resultant amplitude OA at first increasing, then decreasing to a second zero, after the manner shown in a, b, c, and d of Fig. 152. The second zero in intensity is reached when the arc forms exactly two circles, for which condition the path difference between the edge disturbances must be two wave-lengths.

If θ_2 is the angular distance of the second minimum from F, it will be evident that

$$\sin \theta_2 = \frac{2\lambda}{a}.$$

Generally, then, the diffraction pattern for an aperture placed before the objective of a telescope consists of a bright centre, bordered on either side with regions of zero intensity, between which of course there are other maxima. Thus OA, in c of Fig. 152, represents the amplitude of the maximum between θ_1 and θ_2.

The curve showing the changes in intensity for this pattern is given in Fig. 153, while an actual photograph of the appearance is reproduced in Fig. 6, Plate V. The pattern may easily be observed if the student will look at a single filament incandescent lamp a

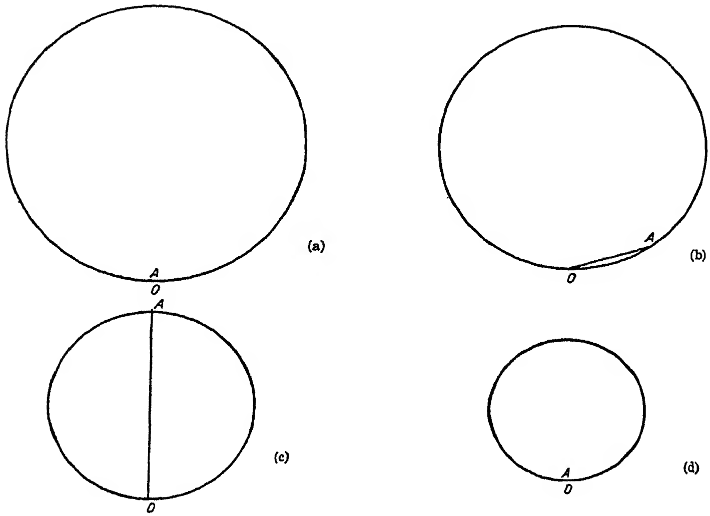


FIG. 152.

few metres away, through the narrow aperture formed by holding two fingers in front of an eye.

It will be noticed that the intensity of the central bright band is much greater than that of the maxima on either side of it. The reason for this follows at once from the explanation which has been given of the whole phenomenon. For, if the aperture is subdivided into n elemental sources, each contributing an amplitude a at F, then the resultant amplitude at F is na. Now, after we pass the first minimum at θ_1, we reach the first maximum on one side when the figure forms (approximately) a circle and a half, as shown in c, Fig. 152. Since, however, there are still n elemental sources and since the amplitude contributed by each

will differ very slightly from a (unless the obliquity is great, which is not usually the case), it follows that the total length of the arc must always differ very slightly from na. OA, then, in c, Fig. 152,

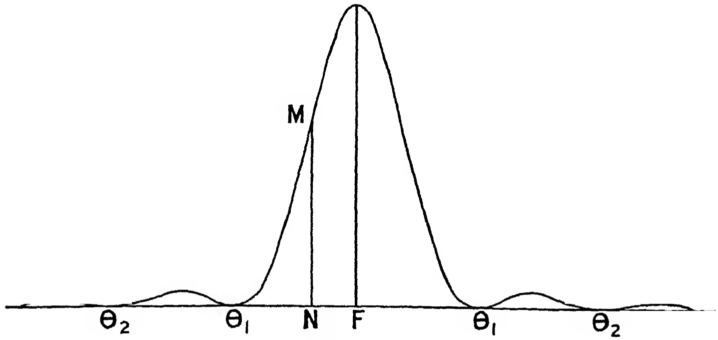


FIG. 153.

representing the amplitude of the first maximum to one side, is the diameter of a circle, three halves of whose circumference is equal to na.

$$\frac{3\pi}{2} OA = na$$

or
$$OA = \frac{2}{3\pi} na.$$

Hence,

$$\frac{\text{amplitude of central maximum}}{\text{amplitude of first maximum}} = \frac{na}{\frac{2na}{3\pi}} = \frac{3\pi}{2}.$$

Therefore, since, as we have proved in section 21, Chapter II, the intensity varies as the square of the amplitude,

$$\frac{\text{intensity of central maximum}}{\text{intensity of first maximum}} = \left(\frac{3\pi}{2}\right)^2,$$

a number which exceeds 22.

We have proved, therefore, that the intensity at the centre of the central maximum is more than twenty times greater than that at the centre of the first bright band on either side, and

obviously still greater than that at the centre of the other maxima. Most of the light, therefore, is concentrated into the central band whose angular width is $2\theta_1$, where

$$\sin \theta_1 = \frac{\lambda}{\quad}$$

102. Rectilinear Propagation. — When the aperture is many times the wave-length of light, θ_1 is very small — indeed, so small, that with ordinary apertures the diffraction bands on either side are not seen at all. Sharp images are then formed, and the laws of geometrical optics hold, — in other words, rectilinear propagation results.

As a becomes smaller, however, θ_1 becomes greater. Thus when the eye looks at a narrow bright source through an aperture $\frac{1}{10}$ mm. wide, if we take 5×10^{-5} cm. as a mean value for λ, we have

$$\sin \theta_1 = \frac{5 \times 10^{-4}}{0.1} = 5 \times 10^{-3},$$

or

$$\theta_1 = 17'.$$

Again, if plane waves fall on an aperture $\frac{1}{1000}$ mm. wide, we find $\theta_1 = 30°$, which means that there is an unbroken band of emergent light about 60° in width.

103. Limit of Resolution and Resolving Power of a Telescope. — It follows from all this that the laws of geometrical optics can be applied to the formation of images by a telescope only when apertures of fair width are utilized — and even then with some limitation, as we shall presently see. The narrower the aperture, the wider the central band, and the less sharp the image. This means that the ability of a telescope to separate two distant objects close together, that is, what is called its *resolving power*, depends on the aperture of the objective. The question of resolving power is so important that we shall here digress to point out that the phrase is used in two senses: (1) when the purpose, as in the case just mentioned, is to examine two

objects close together; or when the fine structure of an object is observed through a microscope; (2) when one has in mind the ability of an instrument to separate two spectral lines (two wavelengths) close together. Concerning the latter some mention will be made in connection with the diffraction grating. At present we are concerned with the former.

Suppose S_1 and S_2, Fig. 154, represent two small luminous objects, either very distant (two near stars, for example) or in the focal plane of the collimating lens of a spectrometer. Then,

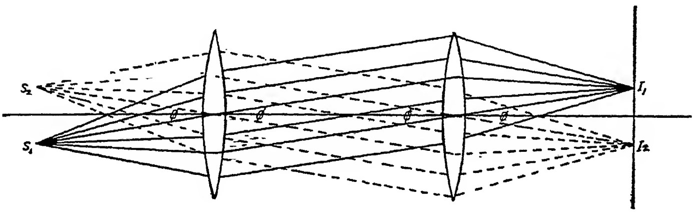


FIG. 154.

if the light emerging from the collimator falls on a telescope, images I_1 and I_2 are formed in the focal plane of the objective. If the aperture of the objective is large, I_1 and I_2 are tolerably sharp, and there is no difficulty in recognizing two distinct images. Suppose, however, that a *rectangular* adjustable aperture is placed before the objective and that its width is gradually decreased. The images I_1 and I_2 then begin to spread out in accordance with Fig. 153, and sooner or later the two central bands will merge into each other. When this is the case, it is not possible to distinguish two images, and therefore to tell that there are two sources. The resolving power of the telescope has become less with the more narrow aperture.

Now, obviously, there must be an aperture of critical width when an observer is just ceasing to be able to distinguish two images or technically, when the *limit of resolution* has been reached. The magnitude of this critical aperture will depend on how close together the two sources are, for evidently the closer I_1 and I_2, the sooner the central bands will merge together. What, then,

is the relation between the separation of the two objects and the corresponding limiting aperture, beyond which it is impossible to tell that there are two?

The answer to this question will be explained with reference to Fig. 155 and Fig. 156. To begin with, if ϕ is the angular separation of the two sources, this is also the angular separation of the *centres* of the two images, *regardless of their total widths.*

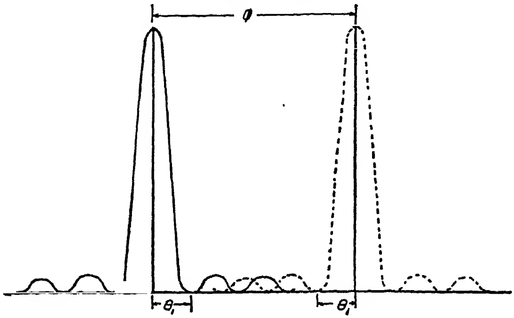


FIG. 155.

When, therefore, ϕ is much greater than θ_1, as represented in Fig. 155, there is no difficulty in seeing two distinct images. As the aperture narrows, θ_1 becomes greater, ϕ remaining constant, and, for a normal eye, the limit of resolution has been reached when θ_1 has increased to such an extent that the first minimum for one image falls on the centre of the central maximum for the other, somewhat as shown in Fig. 156. While different observers may differ slightly in stating just when the two images fuse, it is agreed to adopt as a measure of the limit the criterion just given, that is, that it is reached when θ_1 becomes equal to ϕ.

Conversely, if we are dealing with a fixed aperture, the angular separation between two objects which a normal eye is just ceasing to distinguish as two is given by $\phi = \theta_1$,

or, since $\sin \theta_1 = \dfrac{\lambda}{a}$ and since θ_1 is small, for ordinary apertures by

$$\phi = \frac{\lambda}{a}. \qquad (11.02)$$

This, then, is a measure of the limit of resolution of a telescope with a rectangular aperture of width a. For a circular aperture of diameter a, the same general ideas apply, and the more extended mathematical theory shows that the limit is then defined by

$$\phi = 1.22 \frac{\lambda}{a}. \tag{11.03}$$

An interesting qualitative, and approximately quantitative, illustration of the truth of the above theory is obtained if the slit of a spectrometer is replaced by a piece of fine wire gauze

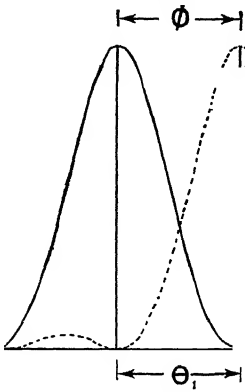


Fig. 156.

illuminated by some monochromatic source such as a sodium flame. With full aperture of the telescope a very sharp image of the pattern of the gauze is seen. If, however, with one set of wires vertical, an adjustable rectangular vertical aperture is placed before the objective, it is found that, starting with a wide opening, when a certain critical width is reached, the vertical wires of the gauze have disappeared altogether. When this is the case, the images of the narrow vertical sources (between each pair of wires) have become so wide that they fuse into one another, in accordance with the explanation already given.

When it is desired, therefore, to detect distant objects which are close together, that is, to increase the resolving power of a

telescope, it is necessary to increase the aperture. For that reason, as well as to increase the amount of light collected, the diameters of the objectives of telescopes used in observatories are very large. At the famous Yerkes Observatory of the University of Chicago, the telescope objective has a diameter of 40″ or approximately 100 cm. With this telescope, therefore, objects with an angular separation as small as $\frac{1.22 \times 5 \times 10^{-5}}{100}$ radian or about one-eighth of a second can be resolved.

Since a normal eye can detect only an angular separation of 1 or 2 minutes, distant objects would not be seen as two unless the magnifying power of the telescope was sufficiently high. Thus, in the case of the Yerkes telescope, taking 1.5 min. as a mean value for the limit of resolution of the eye, a magnifying power of $\frac{1.5 \times 60}{\frac{1}{8}}$, or 720, is necessary.

It will be recalled that in section 54, Chapter V, it was pointed out that high magnifying powers should go along with wide apertures for another reason — in order to keep the diameter of the eye-ring no bigger than that of the pupil of the eye.

103A. Resolving Power of a Microscope. — When the term numerical aperture (N.A.) was introduced in section 56A it was stated that this quantity was of importance chiefly in connection with the resolving power of a microscope. We are now able to discuss this question.

In the previous section we have shown that the smallest *angular* separation of two distant objects which a telescope can resolve is $\frac{1.22\lambda}{a}$, where a is the aperture of the objective. When we use a microscope, the corresponding problem is to find the smallest *linear* separation which two neighboring small objects can have and be seen as two in the instrument. Fundamentally the two problems are the same because in each case the resolving power depends on the effect of diffraction, and in each case the

limit of resolution is reached when two neighboring images overlap to the extent shown in Fig. 156.

Suppose, in Fig. 156A, that K is the position of the centre of the image of the small object R formed by an objective, represented by the lens AB. For reasons discussed in section 110, as we go out from K toward the edge of the image, the intensity of the light falls to zero at a point M for which the path difference $BM - AM = 1.22\lambda$.

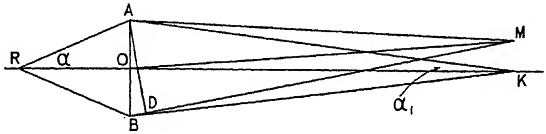


Fig. 156A.

Since AB represents the objective of a microscope, OK is many times greater than RO. Hence, by making $DM = AM$ and joining OM and AD, we can show by exactly the same method as was used in section 79, that, with little error,

$$\frac{BD}{AB} = \frac{KM}{OK}.$$

Therefore, MK, the radius of the central bright portion of the image

$$= \frac{BD \cdot OK}{AB} = \frac{1.22\lambda}{2 \tan \alpha_1},$$

where α_1 = image slope angle AKO.

If we are dealing with two small neighboring objects, the limit of resolution is reached when M represents both the place of zero intensity at the edge of the image of one object and the centre of the image of the other. In other words, MK is also the linear separation of the centres of the two images when the limit of resolution has been reached.

Therefore, d, the linear separation of the two closest objects which can be resolved, is found at once from the relation

$$d = \frac{MK}{\text{initial magnification}}$$
$$= \frac{1.22\lambda}{2m \tan \alpha_1},$$

where m is the symbol previously used for the initial magnification of an objective. From the sine condition we have

$$m = \frac{n \sin \alpha}{n_1 \sin \alpha_1},$$
$$= \frac{\text{N.A.}}{\sin \alpha_1},$$

since in the microscope the image medium is always air. Hence,

$$d = \frac{1.22\lambda \cdot \sin \alpha_1}{2 \tan \alpha_1 \cdot \text{N.A.}}.$$

Since OK is many times greater than AB, α_1 is small, and hence, with little error $\tan \alpha_1 = \sin \alpha_1$, and we have finally

$$d = \frac{1.22\lambda}{2\ \text{N.A.}}. \qquad (11.03a)$$

The resolving power of a microscope may be increased, therefore, by using as short a wave-length and as large a N.A. as possible. The dependence on wave-length is just what we should expect, because the shorter the wave-length the less the diffraction. For visual observation there is not much leeway in the choice of wave-length since we are more or less restricted to the use of white light, for which the brightest part of the spectrum is around 5300 angstroms. In photomicrography, however, there is a very distinct gain in resolving power by using the shortest wave-length possible. For example, with a N.A. of 1.00 some 96,000 * lines to the inch can be resolved if $\lambda = 5300$, whereas,

* It is useful to remember that, in round numbers, some 100,000 lines to the inch are resolvable by a N.A. of 1.00.

if $\lambda = 4000$, about 127,000, or well over 25 per cent. more are resolvable. Photomicrograph equipment has been built in which a still greater gain in R.P. is obtained by the use of the intense radiation emitted by mercury vapour in the neighborhood of $\lambda = 3650$.

When the object medium is air, 0.95 is the highest value of N.A. which is available. With an objective of this N.A., relation (11.03a) shows that we cannot resolve a structure in which details are separated less than approximately $\frac{\lambda}{2}$. By the use of an oil immersion objective, the value of the N.A. may be increased to a maximum of 1.4, with a corresponding gain in resolving power.

Magnifying Power and Resolving Power. — Although the resolving power depends entirely on the objective, the eyepiece plays an important part in visual observation. The details of the real image formed by the objective may be well resolved, but this detail cannot be observed by the eye unless there is sufficient magnification. The point is illustrated by a concrete example.

Suppose a microscope has an objective with N.A. equal to 0.95, which, as we have seen, will resolve small objects $\frac{\lambda}{2}$ apart. Two neighboring objects separated by this distance when viewed directly by the eye at 25 cm. subtend at the eye an angle of about $\frac{5.3 \times 10^{-5}}{2 \times 25}$ or approximately 10^{-6} radian. Now a normal eye can separate neighboring images on the retina only if their angular separation is at least 1 or 2 minutes. Using $1\frac{1}{2}$ minutes as a mean value, we find, since $1\frac{1}{2}$ mins. $= 4.5 \times 10^{-4}$ radian, that the magnifying power necessary to see the resolved images of the two objects $\frac{\lambda}{2}$ apart must be at least $\frac{4.5 \times 10^{-}}{10^{-}}$ or about 450.

In section 56A we have shown that, for a N.A. of 1.00, the normal M.P. is only 200, or possibly 250. In the same section it was pointed out that the use of higher than normal M.P. cuts down the size of the exit-pupil. The full N.A. of the objective is

used, however, and hence the increased M.P. does not involve a loss of resolving power. At this stage we again point out that, if M.P. less than normal are used, the N.A. of the objective is not fully utilized and there is a consequent loss in resolving power.

Microscope Condensers. — Since the objects examined by a microscope are not self-luminous, they must be properly illuminated. The full value of the N.A. of an objective cannot be utilized unless the illuminating cone of light has a sufficiently great angular width. To ensure this, a good microscope is provided with one or more substage *condensers*, that is, with a combination of lenses which collects the illuminating rays of light and throws a highly convergent beam on the small object. The method is illustrated by Fig. 156B, a diagram taken from literature of the Bausch and Lomb Optical Company.

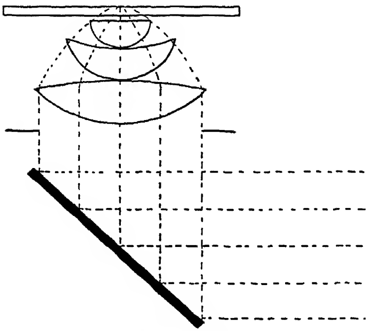


FIG. 156B.

104. Two Narrow Apertures (The Double Slit). — In Chapter VIII we have shown that when a double slit is placed before the objective of a telescope, interference fringes are observed in its focal plane. In the discussion of that problem it was more or less assumed that each of the slits was extremely narrow and that we had to do only with the superposition of two interfering beams. Now, as a matter of fact, when interference fringes are obtained in this way, the width of the slits is often quite comparable with that of the opaque region between them. Although no objection need be made to the statements made in the previous discussion, for a complete understanding of the problem consideration must be given to the diffraction of a single opening.

Suppose, then, that we have two apertures of equal width, AB, CD, Fig. 157, and that we wish the resultant amplitude at

any point P in the focal plane of the objective lens of a telescope. For each aperture separately we can find the resultant amplitude in the manner already described. Thus the resultant for each slit

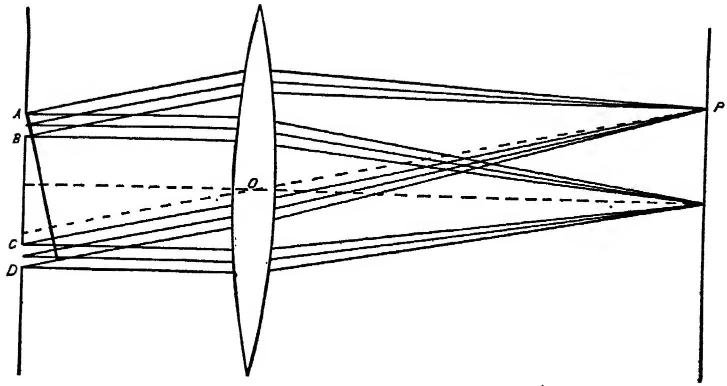


FIG. 157.

alone might be represented by MN in the curve of Fig. 153. When the two slits are acting together, therefore, the final general resultant at P is found by combining two amplitudes, each equal

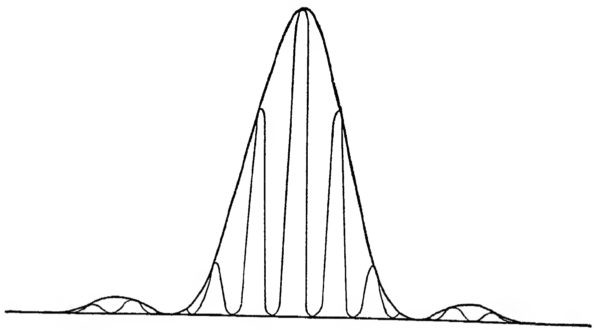


FIG 158.

to MN, but differing in phase by an amount which depends on the position of P, Fig. 157. The problem, therefore, is essentially one of interference.

Since for any particular position of P the path difference between any ray in one slit and the corresponding ray in the next is always the same (for example, the ray from edge A differs in path from the ray from edge C by the same amount that a ray from B differs from a ray from D), the difference in phase between the two resultant disturbances meeting at P can be found from any such pair of rays. Thus, if the central rays differ in path by half a wave-length, the resultant phase difference will be π,

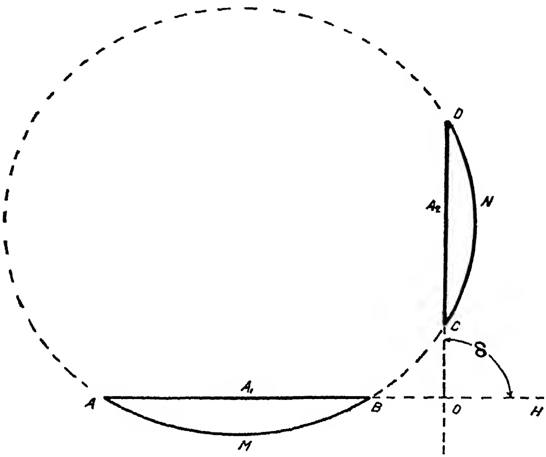


FIG. 159.

and P is dark; if this path difference is a whole wave-length, P is bright, and so on, just as we have already seen in the work on interference. We see now, however, that the interference fringes are superimposed, as it were, on the diffraction pattern of a single slit, the intensity distribution being somewhat as represented in Fig. 158. That this is the case is clearly shown in the photograph reproduced in Fig. 7, Plate V. To obtain the double slit fringes of Fig. 1, Plate IV, a weak source of light and a shorter exposure were used, so that the fringes there seen are only the more intense ones in the central band of the single slit pattern.

The problem of the double slit may also be discussed by making

use of the same graphical method as was used for the single aperture. The solution is indicated by the diagram of Fig. 159, where A_1 represents the amplitude at some point such as P, Fig. 157, due to the first slit, alone; A_2 that due to the second slit, while the angle COH $(= \delta)$ is the difference in phase between A_1 and A_2. It is not difficult to see that this angle is also equal to the difference in phase between the disturbances at P from the central elements of each opening, M and N, or from any other corresponding portions of the two openings.

105. Many Apertures (The Transmission Grating). — While the methods we have been considering may be extended to a consideration of three, four, five or any number of apertures, we shall confine our discussion to the case of a large number of narrow rectangular apertures, alternating with opaque regions. Such an arrangement constitutes a *transmission grating*, although in the gratings in actual use the openings are not quite so sharply defined as we have represented in Fig. 160. The fundamental principles, however, may be arrived at by a consideration of the more or less ideal grating, a portion of which is depicted in the figure.

In the first place, it should be evident that just as in the case of two openings, at any point such as P, Fig. 160, there is a resultant amplitude due to each aperture alone, which is the same for all apertures, at the same point. To find the general final resultant at P, then, we must combine a large number of such amplitudes, between every two of which there is the same difference in phase. Moreover, this difference in phase, just as in the case of the double slit, is that existing between the rays from any two corresponding portions of any two successive openings. When, therefore, the path difference between rays from points such as A and C $(= CM)$, or C and E $(= EN)$, or E and G $(= GL)$ is exactly one wave-length, or two or three, etc., the disturbances at P arising from all apertures are in the same phase, and for such directions, very bright maxima are obtained.

If θ_1 is the angular distance from F of the first bright maximum, θ_2 of the second, or generally θ_n, of the nth, then since $\triangle sPOF$

and ACM (or CEN or EGL) are similar, we have $\angle POF = \angle CAM$ (or $\angle ECN$ or $\angle GEL$), and hence

$$\sin\theta_1 = \frac{CM}{AC},$$

or $$\lambda = AC\sin\theta_1,$$
and similarly $$2\lambda = AC\sin\theta_2,$$
or generally $$n\lambda = AC\sin\theta_n. \qquad (11.04)$$

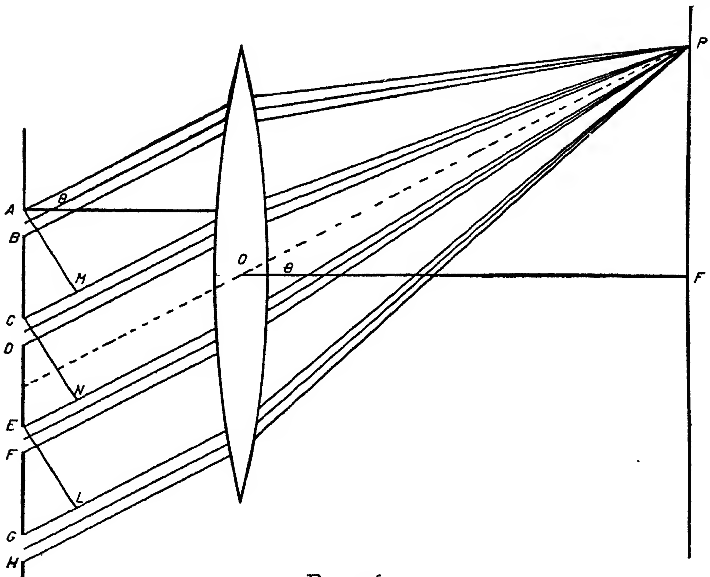


FIG. 160.

AC, the width of one opening and one opaque region is called the *grating element*. We shall designate the magnitude of the element by $a + b$, a for the width of the aperture, b for that of the opaque part. The value of $a + b$ can be found by means of a high-power travelling microscope, or frequently it is provided when the grating is manufactured.

106. Width of Maxima. — When CM or EN or GL, Fig. 160, is equal to an odd number of half wave-lengths, the disturbances

from successive openings differ in phase by π and the resultant is a minimum of zero intensity. It is important for the student to realize, however, that in the ordinary grating with many apertures the intensity does not *gradually* fall from a maximum to a minimum giving broad bright bands separated by dark, but, on the contrary, that there are extremely narrow bright lines separated as a rule by wide dark regions. While the general reason for this has been indicated in the discussion of the Fabry and Perot interferometer

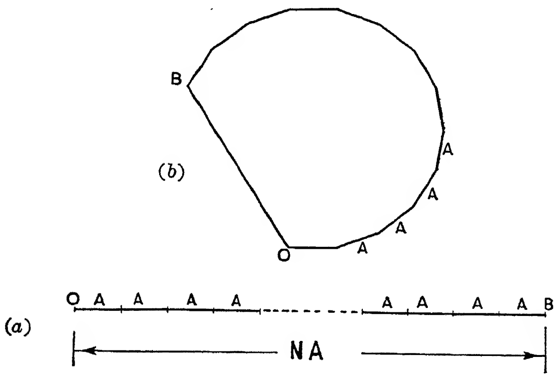


FIG. 161.

fringes, the problem is somewhat different in the case of a grating where we have to do frequently with several thousand fine openings.

Suppose we have N openings and that A is the amplitude due to a single one, at a value of θ where a maximum occurs. The final resultant amplitude at this maximum is then NA and is represented by the line OB, in *a* of Fig. 161. If now we consider points corresponding to values of θ *slightly* greater than that of the maximum, there exists a phase difference between every two successive openings and the general resultant is represented by OB, in *b* of Fig. 161, the total length of the polygon being still NA. (If N is large, as is generally the case, the figure will form an arc of a circle.) Now, applying the same ideas as were used in the case of a single slit, we see that the figure will close, that is, the

total resultant intensity will drop to zero, when the phase difference between the disturbances from the first and the *last* openings has increased by 2π. (More exactly by $\frac{N-1}{N} \cdot 2\pi$, but for very large values of N, the error is slight and does not affect the point under discussion.) The corresponding phase difference between disturbances from the first and the *second* openings, or any two successive ones, is therefore $\frac{2\pi}{N}$, while the path difference for these two disturbances is $\frac{\lambda}{N}$. We have shown, then, that the intensity drops from a maximum to zero when the path difference between disturbances from the first and second openings has changed by $\frac{\lambda}{N}$. Since at the maximum itself, this path difference is either λ or 2λ or $3\lambda \cdots$, it follows that the *change* in the value of θ giving what we may call half the width of the maximum, must be extremely small when N is as large as it is in ordinary gratings. Thus, dealing with the first maximum, if α is one-half its angular width, we have

$$\sin \theta_1 = \frac{\lambda}{a+b}$$

$$\sin (\theta_1 + \alpha) = \frac{\lambda + \frac{\lambda}{N}}{a+b}, \qquad (11.05)$$

from which

$$\frac{\sin (\theta_1 + \alpha)}{\sin \theta_1} = 1 + \frac{1}{N}.$$

For large values of N, therefore (and N may be 10,000 or 20,000 etc.), α must be a very small angle compared with θ_1. Consequently when a transmission grating of many elements is placed on the table of a spectrometer in the path of the parallel beam of monochromatic light leaving the collimator, and the diffracted light is observed in a telescope, one observes in addition to the central image at F, Fig. 162, *narrow* maxima at I_1 and I_1' at an

angular distance of θ_1 from the centre; other maxima (not shown) at I_2 and I_2', at a distance of θ_2 from the centre and possibly still others. These maxima, which are usually called spectral *orders*, are so sharp that they appear as images of the slit of the collimator, just as (but for a very different reason) in the case of a prism. It can be shown by analysis beyond the scope of this book that, in

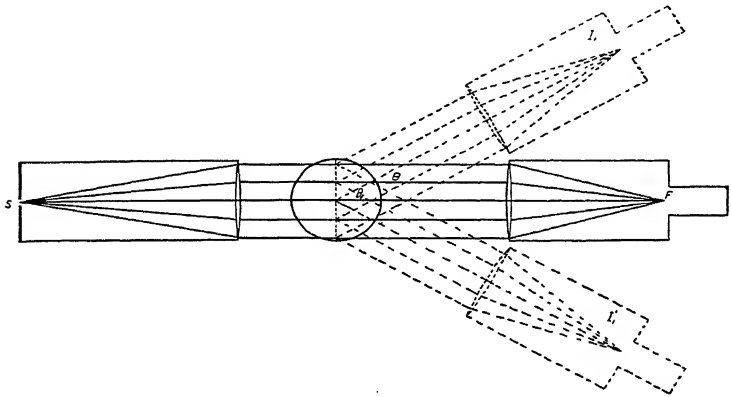


FIG. 162.

between these orders, there are secondary maxima. In the case of an ordinary grating, however, these are so numerous and so faint, that they are not visible.

107. Measurement of Wave-Length. — Since the angles θ_1 and θ_2 can be measured with a fair degree of accuracy by the spectrometer, the wave-length of the light used can be calculated (from relation 11.04) more accurately than by any of the other methods which were discussed under interference.

It follows, also, that when a source which is not monochromatic is used, for each wave-length present the corresponding maximum for any order occurs in a different position. We thus obtain in each order, a spectrum of the source. It will at once be evident from the fundamental relation (11.04) connecting wave-length and angular deviation that all grating spectra, unlike prismatic,

are rational. The relative displacement of corresponding lines is always the same.

108. Number of Orders. — The number of orders which can be observed depends both on the width of a single aperture and on the size of the grating element. The larger the grating element, the smaller the angular separation between orders, and, therefore, the greater the number of orders visible — *provided the diffracted light spreads out over a sufficiently large angle.*

To illustrate — if the grating element = 0.01 cm., and light for which $\lambda = 5 \times 10^{-5}$ cm. is used, we have

$$\sin \theta_1 = \frac{5 \times 10^{-5}}{10^{-2}} = 0.005,$$

or

$$\theta_1 = \text{about } 17'.$$

Similarly θ_2 is about $34'$; θ_3, $52'$; and so on.

Again, if, for the same wave-length, the element = 0.001 cm. we find that $\theta_1 = 2° 52'$; $\theta_2 = 5° 44'$; $\theta_3 = 8° 38'$.

But the angular width over which light is spread out depends on the width of a single opening, the central band (which as we have seen represents almost the whole of the light) having an angular width 2θ, where $\sin \theta = \frac{\lambda}{a}$. In the first example, therefore, if a, the width of the *opening*, is 0.004 cm., light of appreciable intensity would be confined to a band one-half of whose total angular width does not exceed $\sin^{-1} \frac{5 \times 10^{-5}}{0.004}$ or about $43'$. Since, with this grating we have found the orders to be separated about $17'$, it follows that only two of appreciable intensity could be observed on either side of the central image.

If, however, *for the same grating element*, each opening had a width of 0.001 cm. then one-half of the angular width of the central band would be approximately $\sin^{-1} \frac{5 \times 10^{-5}}{0.001}$ or $2° 54'$. In this case, then, as many as $\frac{2° 54'}{17'}$ or 17 on each side are theoretically possible.

In the laboratory, transmission gratings are frequently used which have about 14,000 grating elements or "lines" to 1 inch. With such a grating, we have, for $\lambda = 5 \times 10^{-5}$ cm.,

$$\sin \theta_1 = \frac{5 \times 10^{-5}}{2.54} \times 14{,}000 = 0.276$$

or $\theta_1 = 16° 2'$.

Also $\theta_2 = 33° 32'$,

and $\theta_3 = 57°$.

We see, then, that even if the openings of this type of grating were sufficiently small to spread out the light the whole 90° on either side of the centre, at the most not more than three orders could be observed. Actually two are generally visible with good intensity.

109. Examples of grating spectra can often be observed with the simplest means. If one views a narrow white source of light through half-closed eye-lashes, the pupil of the eye is crossed with a coarse grating, and coloured bands corresponding to successive orders are seen on either side of the central image. The same phenomenon is observed if vision is through a piece of fine wire gauze, with this difference, however, that there are now two sets of diffracted orders, one for the wires running in one direction, the other for those at right angles. Instead of the wire gauze, suitably "ribbed" semi-transparent cloth goods, such as is sometimes found in umbrella covers, may be used. It may not be amiss to point out that the grating was first used to obtain spectra and to measure wave-lengths by Fraunhofer, who in 1821 read a paper on this subject. It is interesting to note that his first gratings, by means of which he made surprisingly accurate measurements, consisted of regularly spaced silver wires of various sizes and at varying distances apart.

110. Dispersion and Resolving Power of a Grating. — In the previous chapter it was shown that the ability of a telescope to separate distant objects close together depends on the aperture

of the objective, and it was there pointed out that we sometimes are interested in another kind of resolving power, namely, the ability of an instrument to separate two wave-lengths close together. Now a grating, as we have just seen, is an excellent means of obtaining the spectrum of a source of light, and therefore in connection with it, as with any instrument which analyzes light into a spectrum, we wish to know two things, (1) the dispersion it produces, (2) the resolving power.

Strictly speaking, as we have already pointed out in Chapter VI, dispersion refers to the change in velocity in a medium with changing wave-length, as a result of which a prism gives a spectrum. The term, however, is also used in the sense with which we are at present concerned, and that is, with reference to the separation of wave-lengths produced by any instrument. When so used (see again section 62, Chapter VI) it is measured by the value of $\dfrac{d\theta}{d\lambda}$.

In the case of a transmission grating, therefore, since for the first order we have

$$\sin \theta_1 = \frac{\lambda}{a + b},$$

by simple differentiation, we obtain

$$\frac{d\theta}{d\lambda} = \frac{1}{(a + b) \cos \theta_1}.$$

More generally, for the nth order, we have

$$\frac{d\theta}{d\lambda} = \frac{n}{(a + b) \cos \theta_n}. \tag{11.06}$$

The dispersion, therefore, increases with the order, and is greater, the smaller the grating element.

Dispersion must not be confused with resolving power, which refers to the degree of closeness two spectral lines can have and still be distinguished as two. The point will perhaps be clear by comparing *a* with *b*, in Fig. 163. In these diagrams, where images of two wave-lengths λ_1 and λ_2 are represented, and where in each MN, the angular separation, is equal, the dispersion is the same for each. It should be evident, however, that with

the wider images as represented in *b*, the two images would fuse into each other much sooner than would be the case with the sharper images of *a*, and that, consequently, the resolving power in the former case would be much less.

To measure the resolving power we make use of a criterion similar to that given in section 103. According to this the

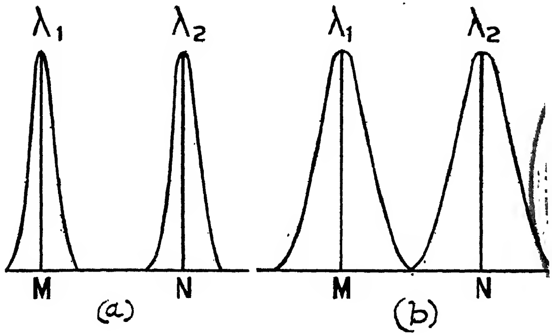


FIG. 163.

images of two spectral lines can just be distinguished as two when the centre of one falls on the edge of zero intensity of the other. Suppose, now, that this critical state is reached when λ_1 and λ_2 differ by $\triangle\lambda$. Then, from relation (11.06), the corresponding angular separation $\triangle\theta$ of the centres of the two lines, in the first order, is given by

$$\triangle\theta = \frac{\triangle\lambda}{(a+b)\cos\theta}.$$

When $\triangle\theta$ becomes equal to α (the angular distance from the centre of either image to the edge of zero intensity), as in Fig. 164, the limit of resolution has been reached. Now α we can find from relations (11.04) and (11.05).

Thus, we have $$(a+b)\sin\theta_1 = \lambda,$$

and $$(a+b)\sin(\theta_1+\alpha) = \lambda + \frac{\lambda}{N},$$

or expanding,

$$(a+b)\sin\theta_1\cos\alpha + (a+b)\cos\theta_1\sin\alpha = \lambda + \frac{\lambda}{N}.$$

Since α is a very small angle, with slight error we may put $\cos\alpha = 1$ and $\sin\alpha = \alpha$, and therefore the last relation becomes

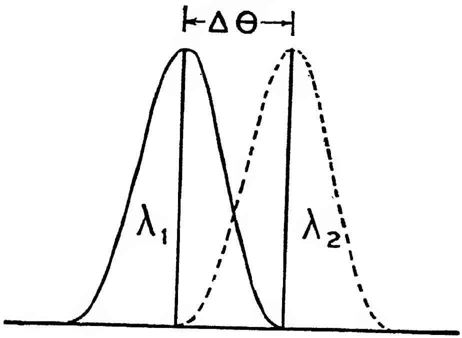
$$(a + b)\sin\theta_1 + (a + b)\cos\theta_1\cdot\alpha \qquad \lambda + \frac{\lambda}{N},$$

or

$$(a + b)\cos\theta_1\cdot\alpha = \frac{\Lambda}{N},$$

or

$$\alpha = \frac{\quad}{N(a + b)\cos\theta_1}. \qquad (11.07)$$



Hence we have the following condition which obtains when $\Delta\lambda$ is the smallest difference in two wave-lengths which can be detected.

$$\frac{\Delta\lambda}{(a + b)\cos\theta} = \frac{\lambda}{N(a + b)\cos\theta},$$

or

$$\frac{\Delta\lambda}{\lambda} = \frac{1}{N}.$$

The magnitude of $\dfrac{\lambda}{\Delta\lambda}$ is taken as a measure of the resolving power.

Thus, if 10,000 elements of a grating were utilized, it follows that in the neighbourhood of 6000 A, one could just distinguish, in the first order, two wave-lengths as close as $\dfrac{\lambda}{N}$ or $\dfrac{6000}{10{,}000}$ or 0.6 A.

When we deal with the nth order, the same type of proof leads to the more general expression

$$\text{resolving power} = \frac{\Lambda}{\Delta\lambda} = n \cdot N. \tag{11.08}$$

The student who finds difficulty in following the mathematical development should be able to see generally (from section 106)

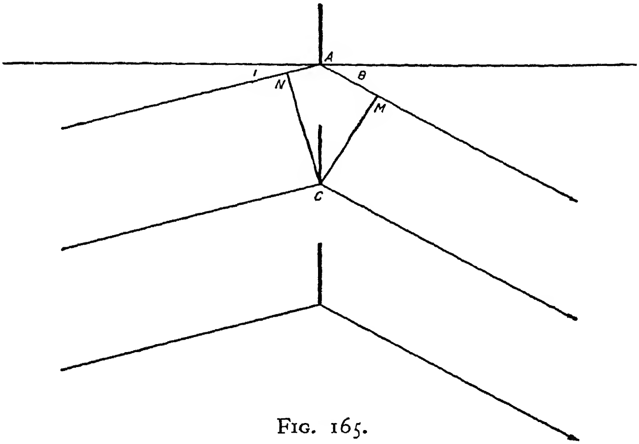


FIG. 165.

that the greater the number of elements, the finer the spectral orders, and, therefore, the closer two wave-lengths can be, and still be separated.

111. Hitherto we have assumed that the incident beam of light strikes a grating normally. This, however, need not be the case, for, if the beam is incident, as in Fig. 165, making an angle i with the normal to the grating, there will be a diffracted beam in any direction making an angle θ with the normal. In such a case the path difference between rays from corresponding points such as A and C, is now $AN + AM$, or $(a + b)(\sin i + \sin \theta)$. The same general considerations as have already been given, still apply and we have

$$\lambda = (a + b)(\sin i + \sin \theta_1) \text{ at the first order}$$

and $$n\lambda = (a + b)(\sin i + \sin \theta_n) \text{ at the } n\text{th order.} \tag{11.09}$$

If the diffracted beam is on the side of the normal opposite to that of the incident beam, we have

$$n\lambda = (a + b)(\sin i - \sin \theta_n). \qquad (11.10)$$

In the case represented by Fig. 165 the nth order is deviated from the incident beam an amount equal to $i + \theta_n$. The following proof shows that the deviation is a minimum when the grating is so placed that $i = \theta_n$.

By elementary trigonometry relation (11.09) may be written

$$n\lambda = 2(a + b) \sin {i +} \cdot \cos {i - \theta_n}$$

or

$$\sin {i + \theta_n} \qquad \frac{n\lambda}{2(a + b) \cos {i -}}$$

From this equation we see that the value of $\sin \frac{i + \theta_n}{2}$, and, therefore, of $i + \theta_n$, is a minimum when

$$\cos \frac{i - \theta_n}{2} = 1, \qquad \text{or when} \qquad i = \theta_n.$$

112. Reflection Grating. — Spectra are also obtained when diffracted reflected light is utilized. Here again the same general considerations hold and it is not difficult to prove that there are diffraction orders subject to the relation

$$n\lambda = (a + b)(\sin i \pm \sin \theta), \qquad (11.11)$$

where i is the angle the incident beam makes with the normal, θ that of the reflected beam. While reflection spectra may be observed with the ordinary transmission gratings so frequently used in the laboratory, good plane reflection gratings are made by carefully ruling, with a diamond point and an accurate dividing engine, equidistant parallel lines on a polished plane surface of speculum metal. The gratings used for laboratory instruction purposes are invariably replicas of a good grating made by taking a celluloid cast. It is possible to convert these into reflection gratings by coating them with a thin layer of metal. For details regarding this, and the actual construction of gratings, the student should consult "Baly's Spectroscopy," Vol. I.

113. Concave Grating. — By far the most important grating is one whose surface, *AMNC*, Fig. 166, is spherical and ruled with lines which are the projection of equidistant parallel lines along the plane surface represented by *AC*. Such a grating may be used to give extremely sharp spectral lines without the use of any lenses, provided certain conditions are fulfilled. Suppose *R* is the centre of curvature of the sphere of which the grating surface is a part and that *S* and *P* are two points lying on a circle having *r* the radius of curvature of the grating as diameter. Then

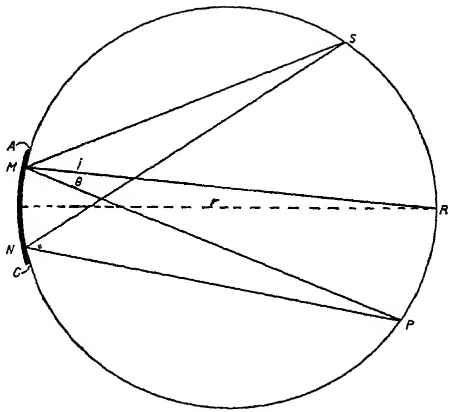


Fig. 166.

mathematical analysis shows that the difference in the paths *S* to *P* between corresponding rays from two successive elements at *M*, is the same for a pair of elements at any other region *N*, for example. Thus, if this path difference is $n\lambda$ for any pair of corresponding rays, it is so for all, and in that case *P* is a bright order, or a spectral image of this wave-length. If a curved screen were placed so as to coincide with a portion of the circle in the neighbourhood of *P*, on it would be observed the various orders of diffracted spectra all sharply focussed. As gratings with radii of curvature as large as 20 ft. or more are often used, while the actual surface of the grating itself is only a very few inches, it

will be evident that in the figure, the width of the grating is enlarged out of all proportion in comparison with the other dimensions.

The first use of a concave grating was made in 1881, by Rowland (United States), who arranged his apparatus to fulfil the above condition after a method now described by his name. In the Rowland arrangement, the grating G, Fig. 167, the screen P (or eyepiece or photographic plate) are at opposite ends of a rigid arm whose length PG is equal to the radius of curvature

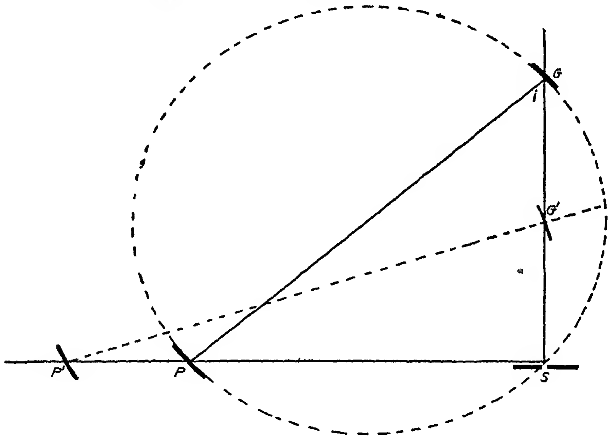


FIG. 167.

of the grating. P and G are on carriages which can run along two rails on the rigid arms GS and PS set at right angles to each other. With this device P, S, and G always lie on the circle whose diameter is PG, and hence, if light emerging from a narrow slit at S falls on the grating and is diffracted, sharply focussed images will be formed in the neighbourhood of P. Moreover, since for the central image, θ (the angle between the diffracted beam and the normal to the grating) is equal to zero, we may write for this image,

$$\begin{aligned} n\lambda &= (a + b) \sin i \\ &= (a + b) \frac{SP}{PG}. \end{aligned} \tag{11.12}$$

It follows that distances along SP are directly proportional to the wave-length in any given order.

To sum up, the importance of the concave grating lies in the fact that by means of it extremely sharp focussed images may be obtained *without the use of lenses*. In the vast work of measuring accurately the multitude of wave-lengths for all elements in various luminous sources it has played a very important part. In modern research it is particularly useful in the measurement of wave-lengths in the extreme ultra-violet region not accessible with ordinary instruments because of the absorption of quartz, of fluorite, and of air itself. For this purpose, gratings with radii of curvature of the order of a metre of two are used, enclosed in a chamber from which the air is exhausted.

114. Scattering by Small Particles. — A beam of light traversing dust-free air under ordinary circumstances cannot be seen even in a darkened room. If, however, dust or smoke is placed in its path, light is reflected from the particles of dust or smoke, and the beam becomes visible. Moreover, if by using coloured glasses, the beam is made approximately monochromatic, red, for example, the light observed in any direction is the same colour. About this there is nothing surprising, for it is just a case of ordinary reflection of light from the particles in its path.

Suppose, however, that a beam of white light traverses a tank of water in which are suspended particles whose diameters are comparable with the mean wave-length of visible light. Such a suspension may conveniently be made by adding to clean water a few drops of milk or a drop or two of an alcoholic solution of gum mastic. It is now observed that the light seen in a direction at right angles to the path of the beam has a distinctly bluish cast. An interesting variation of this simple experiment may be made by having, instead of the gum mastic suspension, a weak solution of hyposulphite of soda to which a few cubic centimetres of hydrochloric acid are added. By this means a precipitate of sulphur particles is obtained, whose number and size are for a short time suitable for studying the light at right angles to a beam

traversing it, as well as the light transmitted and allowed to fall on a screen. In this case the bluish side light becomes very marked, while, as the precipitate increases, the transmitted light takes on a deeper and deeper orange-reddish shade. Evidently an excess of shorter wave-lengths is emitted in a direction at right angles to the beam, while the transmitted beam has a corresponding excess of the less refrangible red and yellow wave-lengths.

This phenomenon is known as *scattering*, and occurs when a train of waves encounters particles whose dimensions are comparable with the lengths of the waves. Because of their size, such particles cannot give rise to reflected wave-fronts but act as independent centres from which disturbances spread out in all directions. The intensity of the scattered light can be shown both experimentally and theoretically to vary inversely as the fourth power of the wave-length; hence the scattered light in the case we have been considering must contain an excess of blue and violet, the transmitted, an excess of the less refrangible rays. Scattering is responsible for certain colour phenomena seen in nature, such, for example, as the blue of the sky, the reddish colour of a setting sun, and the colour of many minerals. Rising smoke seen against a dark background is bluish because there is little or no transmitted light, whereas the scattered light received by the eye is relatively intense. On the other hand against a clear sky or a bright cloud the smoke through which light is passing (from the bright background) may appear brown or reddish.

115. The Blue of the Sky. — In considering this phenomenon it is natural to ask one or two questions. In the first place, why is the sky blue and not violet? To answer this question it is necessary to realize that when an observer at A, Fig. 168, is receiving scattered light from a portion of the blue sky B remote from the sun S, this scattered light must itself traverse a stretch of the atmosphere represented by BA. From this light, therefore, which originally contained an excess of green, blue, and violet, the latter being most intense, there is re-scattering in the direction of the arrows, an excess of the shortest wave-lengths being removed

in this way. The resultant light which finally reaches the eye has a maximum intensity in the blue region.

Again, it may be asked, what are the scattering particles? While extremely fine particles of dust are undoubtedly present in the low layers of the atmosphere, the brilliant blue of the sky, seen over sea as well as land, cannot be ascribed entirely to that cause. Can it be, then, that the cause is found in the scattering

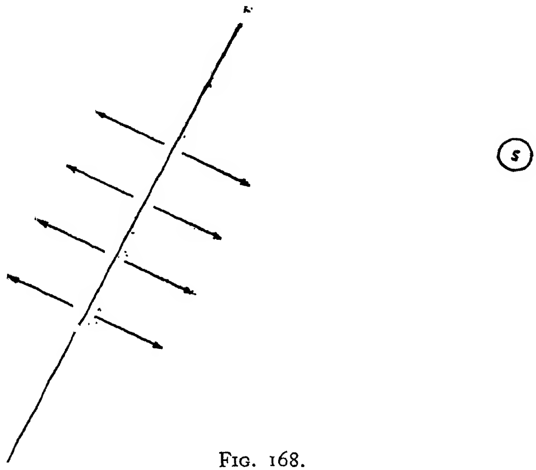


FIG. 168.

of molecules of gases or of water vapour? For some time the answer to that question was a debatable matter, but finally was given by the research work of such men as Lord Rayleigh, who showed by direct experiment that it is possible to observe scattering from molecules of a gas. There can be little doubt then, that scattering from molecules is to a considerable extent responsible for the blue of the sky.

In regard to the red colour of a setting sun, fine particles in suspension no doubt play a more important part, for, in this case, the light travels approximately horizontally through a long layer of air near the surface of the earth. In such a layer fine particles of foreign matter are to be expected.

Problems

1. Briefly explain the difference between light and sound waves in casting shadows.

2. Plane waves, for which $\lambda = 5 \times 10^{-5}$ cm., fall on an aperture, and the resulting diffraction pattern is observed in an eyepiece 100 cm. from the aperture. If the distant source of light is a point, calculate the radius of the 100th half-period zone.

3. A zone plate has the radius of the first ring 0.5 mm. If plane waves of light, wave-length = 0.00005 cm., fall on the plate, find where a screen must be placed so that the light is focussed to a bright spot.

4. Monochromatic light of wave-length 5×10^{-5} cm. and with a plane wave-front falls normally on a screen. In the screen is a pin-hole of one millimeter diameter. How far should a second screen, parallel to the first, be placed from it, so that the intensity of light at a point P on the axis of the pin-hole is a maximum?

5. A circular opening of radius 0.6 mm. is illuminated by plane waves of monochromatic light and the emergent light allowed to fall on the screen. As this screen is gradually brought from a distance toward the opening it is found that when the distance between opening and screen is 30 cm. a dark spot appears at the centre of the patch of light. Find the wave-length of the light.

6. (*a*) What single characteristic of light accounts for the difference between sound and light, in so far as rectilinear propagation is concerned?

(*b*) Find the radius of a pin-hole in a large opaque sheet such that, when it is illuminated with plane waves of monochromatic light ($\lambda = 6 \times 10^{-5}$ cm.) and the emergent light is received on a screen 100 cm. away, the intensity at the centre of the diffraction pattern is the same as at the same point when the opaque sheet is removed.

7. Light from a narrow distant source falls on a narrow rectangular aperture. If the diffraction pattern is observed on a screen, (1) in the plane of the image formed by the lens placed just behind the aperture, (2) with the lens removed, describe the chief difference observed in the two cases as the width of the rectangular aperture is gradually altered.

8. A narrow distant linear source is examined by means of a telescope whose objective has a focal length of 60 cm. Describe *quantitatively* and explain with diagrams the appearance observed on a screen placed in the focal plane of the objective, when a narrow rectangular aperture of width 0.5 mm. is placed before the objective. Assume $\lambda = 5 \times 10^{-5}$ cm.

9. A linear aperture whose width is 0.002 cm. is placed immediately in front of a lens of focal length 60 cm. If this aperture be illuminated with a beam of parallel rays whose angle of incidence is zero, and whose wave-length is 5×10^{-5} cm. what will be the distance between the centre and the first dark band of the diffraction pattern on a screen placed 60 cm. from the lens?

10. Plane waves of monochromatic light ($\lambda = 6 \times 10^{-5}$ cm.) fall on a narrow aperture, width $\frac{1}{5}$ mm., immediately beyond which is a lens of focal length 30 cm. Discuss quantitatively the light pattern in the focal plane of the telescope.

11. The Yerkes Observatory telescope has an objective lens of diameter 40 inches. Assuming a suitable mean wave-length, estimate the smallest angular separation of two stars which can be resolved by it.

12. A spectrometer has *two* very narrow collimator slits with centres 0.1 mm. apart. The focal length of the collimating lens is 20 cm., that of the objective of the telescope is 30 cm. If the eyepiece of the telescope consists of a single lens of focal length 4 cm., find

(1) The *linear* separation of images in the focal plane of the objective;

(2) The *angular* separation of images observed through the eye lens;

(3) The smallest rectangular aperture of the objective for which the two images just fail to be separated. Assume $\lambda = 5.5 \times 10^{-5}$ cm.

13. The telescope of a spectrometer is focussed on a distant object and the collimator slit is then adjusted so that a sharp image of the slit is seen in the telescope.

(*a*) Make a diagram showing the path of a complete bundle of rays for a narrow slit.

(*b*) If a double slit with openings 0.5 mm. apart is now placed on a table of the spectrometer, explain, again by the aid of a careful diagram, why interference fringes are seen.

(*c*) If the mean wave-length of the light used is 5.5×10^{-5} cm., find the angular separation of these fringes.

(*d*) If one of the double slits is completely covered, describe the appearance observed in the telescope.

14. Discuss the importance of diffraction in the construction of a telescope.

15. In a telescope whose objective has a focal length of 50 cm. the eyepiece is replaced by a ground glass screen on which is focussed a small distant source of monochromatic light ($\lambda = 5 \times 10^{-5}$ cm.). If an adjustable rectangular aperture is placed before the objective, find for what width (of this aperture) a continuous central band of light 4 cm. long is seen on the ground glass.

16. A brass plate has two vertical narrow openings cut in it, with centres 2 mm. apart. A sodium flame is placed directly behind these slits, and a telescope 3 metres away is focussed on the slits. A rectangular aperture is then placed immediately in front of the objective of the telescope and its width gradually lessened until it is no longer possible to see two images. What is the width of the aperture when this occurs?

17. A narrow vertical distant incandescent filament is viewed by a naked eye through a single narrow vertical aperture of width 0.5 mm.

(1) Describe and explain the appearance observed.

(2) Find a mean value for the angular width of the central band observed, taking $\lambda = 0.00005$ cm.

18. (*a*) A converging lens forms a real image of a distant object on a screen in its focal plane. Discuss the effect of the aperture of the lens on, (i) brightness, (ii) sharpness, (iii) magnification, of image.

19. If you look through a piece of fine gauze (40 wires to 1 cm.) at a narrow *red* source of light ($\lambda = 6 \times 10^{-5}$ cm.) placed 4 metres from the gauze, what will be the linear separation between the central and the first diffracted image?

20. Plane waves of monochromatic light ($\lambda = 0.00005$ cm.) fall on a piece of fine wire guaze with one set of wires vertical. The emergent light is focussed on a screen by a lens of focal length 100 cm. placed immediately behind the wire grating. If there are 5 wires of the grating to 1 mm. describe quantitatively the appearance on the screen.

21. A glass grating is ruled with 4250 lines to the centimetre. When plane waves of yellow light strike this grating at perpendicular incidence, it is observed that the spectrum of the second order is deviated (diffracted) through an angle of 30°. What is the wave-length of the yellow light?

22. A source of light omits two colours. When examined through a grating, it is found that the spectrum of the 4th order for one colour is seen in the same direction as that of the 5th order for the other colour. What is the ratio of the wave-length for these two colours?

23. In what respects does the spectrum produced by a diffraction grating differ from a prismatic spectrum?

24. An observer, looking through a piece of fine wire gauze, at a yellow flame 3 metres from the gauze, notes several images of the flame on either side of it. By the method of parallax the distance apart of the two images adjacent to the flame (one on either side) is found to be 1.8 cm. Find the wave-length of the yellow light. Size of gauze = 5 wires per mm.

25. A grating is made with 1000 lines to the centimetre, and with openings equal to one-third of the total element. Find, for $\lambda = 5 \times 10^{-5}$ cm.,

(*a*) if only *one* opening is used, approximately the angular distance the light continuously spreads out;

(*b*) the linear distance between the centre and the first order spectrum on a screen placed in the focal plane of a converging lens of focal length 60 cm. Parallel rays fall on the grating.

26. The angular deviation of the fourth order (diffraction grating) spectrum of a certain colour is found to be 30°. If the grating contains 212 lines to a mm. what is the wave-length of the colour?

27. Why are the maxima (bright orders) for a diffraction grating of many openings (1) narrower, (2) farther apart, than the maxima obtained with two slits such as used in an ordinary double slit?

28. If a normal eye cannot separate two objects (or images) subtending an angle of less than 1.5 minutes, find approximately the magnifying power that is necessary to make use of the full resolving power of a telescope, whose objective has a diameter of 50 cm. (Assume a mean wave-length of 5.5×10^{-5} cm.)

29. Plane waves of monochromatic light ($\lambda = 4.5 \times 10^{-5}$ cm.) fall normally on a screen A which is opaque except for a circular hole of radius 0.6 mm. and a concentric annular ring whose inner and outer radii are $0.6\sqrt{2}$ mm. and $0.6\sqrt{3}$ mm. Prove that the *intensity* of the light at the centre of the diffraction pattern received on a screen B placed 80 cm. from screen A is approximately 16 times the intensity at the same point if screen A is removed altogether.

30. A series of concentric circles with radii in the ratio of 1; $\sqrt{2}$; $\sqrt{3}$; $\sqrt{4}$; $\sqrt{5}$; etc. are drawn on a white sheet of paper and every *second* one of the annular areas thus created is densely blackened. The screen is photographed on a small scale so that a corresponding series of alternately transparent and opaque areas are made on a photographic plate or film, with the radius of the first circle equal to 0.8 mm. If parallel rays of monochromatic light of wave-length 6.4×10^{-5} cm. are allowed to fall normally on this photographic plate, and the emergent light is received on a screen 100 cm. away, prove that the intensity of light at the centre of the pattern on the screen is many times greater than the intensity at the same point if the photographic plate is removed altogether.

31. Plane waves of monochromatic light of wave-length 6×10^{-5} cm. fall normally on a grating consisting of a number of narrow slits, width of each 0.01 mm., separated by opaque regions of width 0.3 mm. The emergent light is focussed on a screen by means of a converging lens of focal length 60 cm. Find (i) the total width of the central unbroken band of light arising from a single slit; (ii) the position and number of the spectral orders (principal maxima) due to the whole grating.

32. The objective lens of a large astronomical telescope has a diameter of 60 cm. and a focal length of 900 cm. Find the focal length of an eyepiece such that an eye which is just able normally to distinguish two images 2′ apart will be able to distinguish through the telescope two small distant objects so close that they are just within the limit of resolution of this instrument. Assume a *rectangular* aperture for the objective and a mean $\lambda = 5.5 \times 10^{-5}$ cm.

33. A small object has a fine structure of 35,000 lines to the inch. Find (i) the smallest N.A. the objective of a microscope can have in order to resolve this structure; and (ii) the minimum M.P. necessary to see the resolved structure.

Take $\lambda = 5.3 \times 10^{-5}$ cm. and assume that the smallest angular separation observable by the eye is $1\frac{1}{2}$ minutes.

34. (*a*) Find how many lines to the inch can be resolved by a microscope whose objective has a N.A. of 0.35.

(*b*) If the initial magnification of this objective is 21, find the M.P. of the eyepiece necessary in order to see the resolved structure.

35. What single factor accounts for the fact that light, for all practical purposes, travels in straight lines, whereas sound does not? Discuss fully.

36. Explain why sound entering a room through an open window can be heard almost equally well anywhere in the room, whereas direct sunlight throws on the floor a sharp image of the window aperture.

CHAPTER XII

DOUBLE REFRACTION

116. In the chapters on Interference and Diffraction phenomena have been described, the explanation of which is possible only in terms of some form of wave theory. We have seen how such phenomena provide us with means of measuring wavelengths of light, but, thus far, nothing has been said regarding the nature of the vibrations. Are they transverse or longitudinal, or a combination of both? In sound we know that we have to do with longitudinal vibrations; what about light? The answer to that question will be found by study of the work of the next few chapters, in which for the first time we consider the propagation of light through transparent crystalline substances. Hitherto we have considered only isotropic substances, those for which physical properties are independent of the direction. But if one takes a piece of a substance like crystalline quartz, it is not difficult to show that heat, for example, travels at unequal rates in different directions. In general, for crystals, physical properties do vary with the direction. It is not surprising, therefore, to find that the passage of light through transparent crystals is subject to laws which, in some important respects, differ from those applicable to isotropic substances.

117. The Fact of Double Refraction. — To obtain our introductory ideas we shall make use of *Iceland spar* (calcium carbonate), a substance whose fundamental crystal form is the rhombohedron. In this crystal form, represented by Fig. 169, it is to be noted that at certain corners such as A, the edges form both obtuse and acute angles, of magnitude $101°55'$ and $78°5'$, while at the corners B and D the three edges form obtuse angles only. A line drawn through B making equal angles with each of

the three edges, or any line parallel to this, defines the *optic axis* (O.A.) of this crystal. Note, then, that the optic axis is a *direction*, and moreover, it is not defined by the line joining B and D except for one particular length of the crystal.

In the latter half of the seventeenth century it was shown by Bartholinus of Denmark that Iceland spar had the property of refracting a beam of light in two different directions. The phenomenon may be shown by the following simple experiment. If a narrow beam of light, such as LM, Fig. 170, falls *normally* on one of the faces of such a crystal, *two* spots of light, O and E, are observed on a screen placed in the path of the emerging light. Moreover, if the crystal is rotated about an axis parallel to the beam, the spot O remains stationary, while the spot E rotates about it. A single beam, therefore, has been *doubly* refracted, that is, has given rise to two beams, one of which, at least in this simple experiment, corresponds to the single beam which would have been transmitted had a substance like glass been

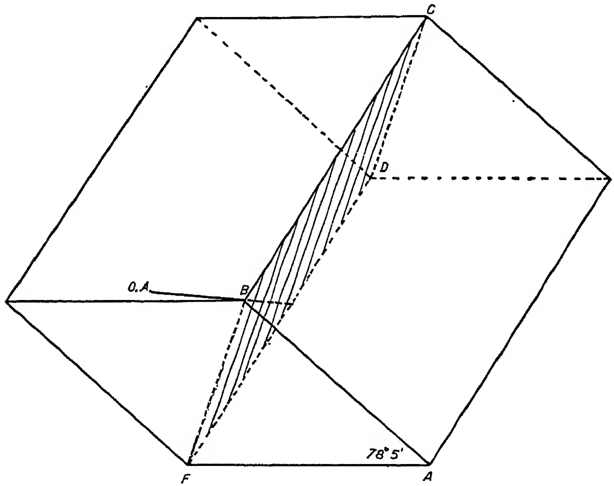


Fig. 169.

used. This beam is called the ordinary (O) beam, the other, the extraordinary (E).

If the beam of light is incident in a direction other than normal, and measurements are made of angles of incidence and of refraction, it is found that the sine law holds for the O beam, whereas for the E beam, the value of $\frac{\sin i}{\sin r}$ varies with the angle of incidence. Now it will be recalled that in section 27 it was proved that, for ordinary media, the sine law holds because $\frac{\sin i}{\sin r}$ is equal to the ratio of the velocity of light in air to that in the denser

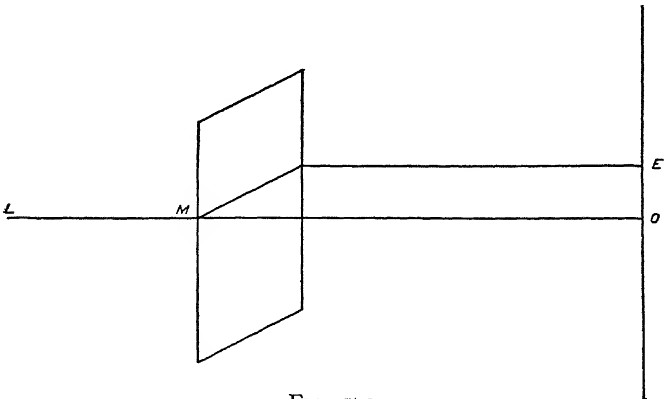


FIG. 170.

medium. Its constancy, therefore, regardless of the direction of the incident light, is a proof that the velocity of light is the same in all directions. Conversely, the lack of constancy of this ratio when we have to do with the E beam is a proof that the velocity of this component in a crystal is different in different directions.

If the position of the crystal with reference to the incident beam is varied, in general two beams are always obtained. When, however, a crystal is used which has been cut so that the beam can be sent in a direction parallel to the optic axis, only one spot

is obtained. We may conclude, then, that in this direction the velocities of the two beams are equal. (See also end of section 119.)

118. Uniaxal and Biaxal Crystals. — Double refraction can be observed in many crystalline substances, such as Iceland spar, quartz, ice, tourmaline, apatite, nitrate of soda, borax, mica, selenite, topaz, and aragonite. An examination of such substances shows that they may be divided into two groups. In the first, called *biaxal*, there are two directions along which the two disturbances travel with the same velocity; in the second, called

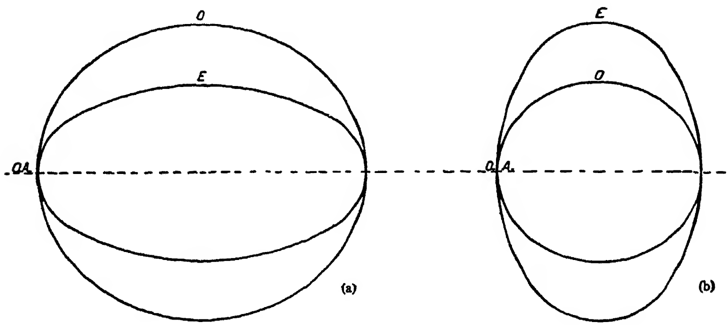


FIG. 171.

uniaxal, there is only one such direction. Of those given above, the last five are biaxal, the remainder uniaxal. While the general explanation of the phenomenon of double refraction shows that the latter is just a special case of the former, in the work of this text we shall confine our attention to uniaxal crystals, of which Iceland spar and quartz will chiefly interest us.

119. Huygens' Explanation of Double Refraction. — In considering the explanation of the phenomenon for uniaxal crystals, enunciated first by Huygens and subsequently verified by experiment, it is well for the student to note again the simple facts.

(*a*) There are actually two beams traversing the crystal,

therefore two wavelets must spread out from a particle which has been disturbed.

(*b*) The velocity of the *O* beam is the same in all directions, therefore the corresponding wavelet must be spherical.

(*c*) The velocity of the *E* beam varies with the direction, but, as already noted, may reasonably be assumed to be the same as that of the *O* beam along the optic axis. Now Huygens assumed that the *E* wavelet was spheroidal, or an ellipsoid of revolution,

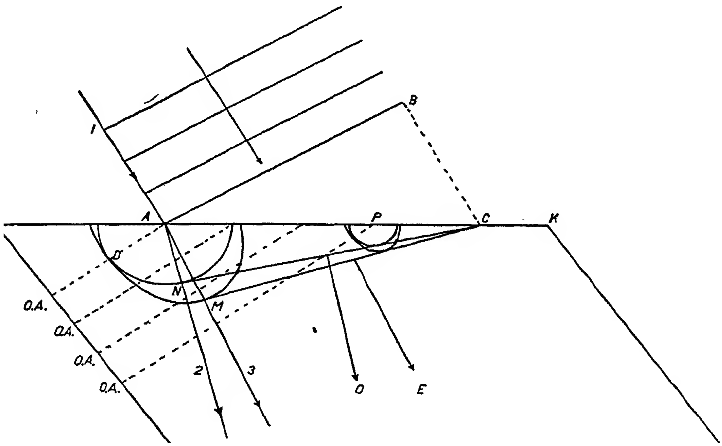


FIG. 172.

the *O* and the *E* wavelets touching at the places where they were intersected by the optic axis. The *E* wavelet, therefore, is formed by the revolution of an ellipse about a diameter coinciding with the optic axis. A cross-section of the two wavelets is then represented by such diagrams as *a* and *b*, Fig. 171. In *a* it will be observed that the *O* wavelet is represented as entirely outside the *E*, whereas in *b* the converse is the case. That both conditions are possible, we shall see presently.

The importance, as well as the use of Huygens' construction, will be evident from a consideration of the following illustrations, in all of which cross-sectional diagrams alone are used.

(*a*) *Plane waves are incident obliquely on a piece of doubly refracting crystal so placed that the O.A. is in the plane of incidence, but subject to no other restriction. Use Huygens' construction to show the existence of O and E plane waves traversing the crystal.*

In Fig. 172, let *AB* be a plane wave-front incident on the plane face *LK* of a crystal whose optic axis is represented by the dotted lines marked O.A. To obtain the refracted wave-fronts we use exactly the same method as was employed in section 27 when ordinary refraction was under consideration. Thus, by the time

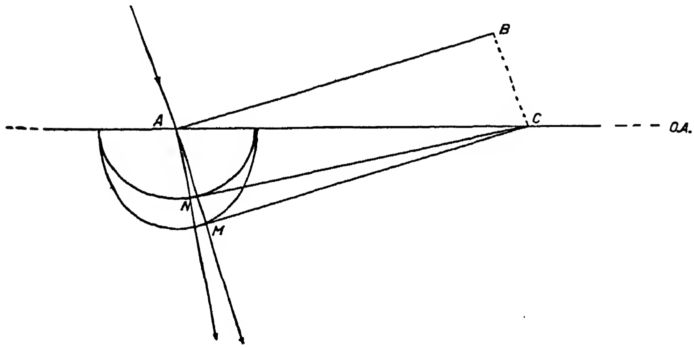


FIG. 173.

the edge *B* of the incident wave-front has reached the point *C* on the surface of the crystal, the *O* wavelet originating at *A* has reached the position *DN*, while the *E* wavelet originating at the same point, has reached *DM*. By drawing the position of a number of other wavelets originating from points such as *P* between *A* and *C*, we finally obtain: (1) a plane wave-front *CN* tangent to all the *O* wavelets, (2) a plane wave-front *CM* tangent to all the *E* wavelets. There are therefore two sets of refracted wave-fronts travelling in different directions. It is to be noted further that, if we consider a ray such as 1, incident at *A*, the corresponding *O* ray 2 is, as in the case of an isotropic substance, perpendicu-

lar to the wave-front *NC*, whereas *the refracted E ray is in general not perpendicular to the wave-front MC.*

(*b*) *When conditions are similar to case* (*a*), *except that the optic axis is not only in the plane of incidence but also parallel to the surface of the crystal.*

No detailed explanation should be necessary for the cross-sectional diagram of this case shown in Fig. 173.

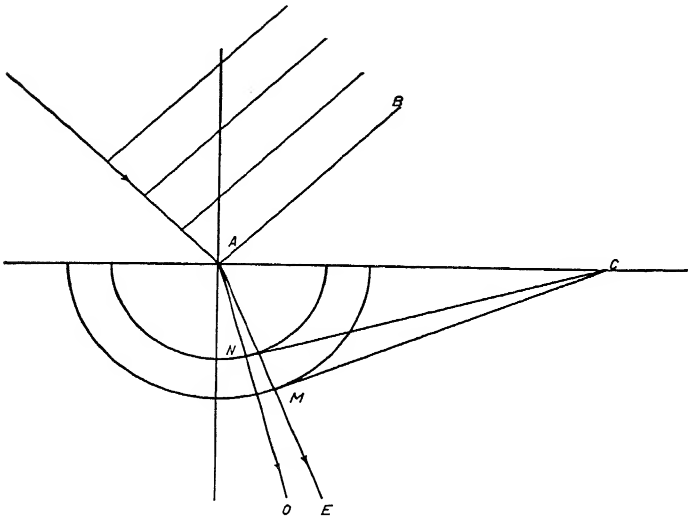


FIG. 174.

(*c*) *Fig. 174 is the diagram for the case when the O.A. is perpendicular to the plane of incidence and parallel to the surface of the crystal.* The student should note this case particularly, for, the O.A. now being perpendicular to the plane of the paper, the sections for both *O* and *E* wavelets are circular. Moreover, they are circles regardless of the inclination of the original wave-front. It follows, therefore, that, provided one works in this plane, the velocity of the *E* disturbance is the same in all directions, and that the sine law holds for both *O* and *E* beams — again, however,

only in this plane. That being the case, we may now agree on the following definition of the extraordinary index of refraction, which we shall represent by n_E.

$$n_E = \frac{\text{velocity of light in air}}{\text{velocity of the } E \text{ beam in a plane perpendicular to the optic axis.}}$$

Cases when plane waves fall normally on the face of a crystal are shown in Fig. 175, where the O.A. is in the plane in which the cross-

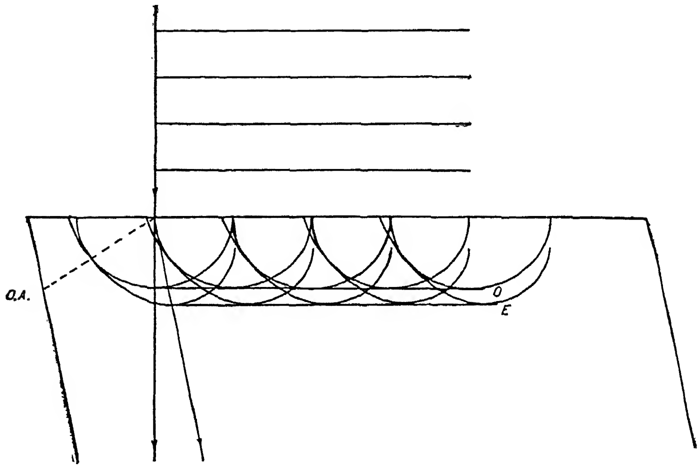


FIG. 175.

section of the wavelets is taken but not otherwise restricted; Fig. 176, where the O.A. is in this plane and parallel to the surface of the crystal; Fig. 177 where the O.A. is perpendicular to this plane but parallel to the surface; and Fig. 178, where the O.A. is perpendicular to the surface of the crystal. Figures 176 and 177, which should be compared with Fig. 178, show that, when plane waves fall normally on crystals cut so that the beam travels *at right angles* to the O.A. there is no separation into two beams, although the two travel at different rates. (Note again end of section 117.)

120. Determination of Refractive Indices. — If a piece of Iceland spar is cut in the form of a prism and placed on the table of a spectrometer, two images are observed, and the angle of minimum deviation can be measured for each. Moreover, if the crystal is so cut that the O.A. is parallel to the refracting edge of the prism, the plane of incidence is perpendicular to the optic axis, and hence the usual calculation by the use of relation (6.03), gives both n_O and n_E.

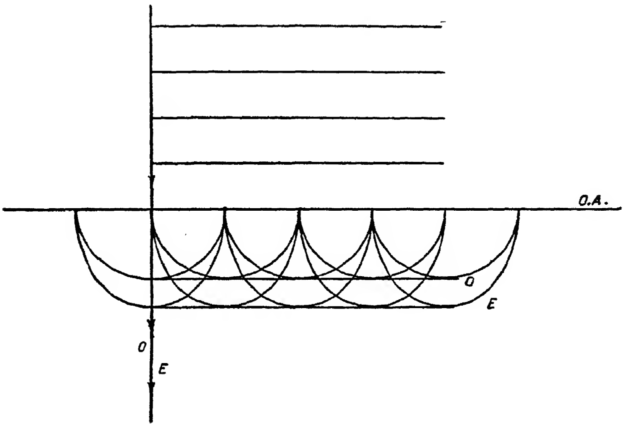


FIG. 176.

In Table IX, values of indices obtained in this or in other ways, are given for a few uniaxal crystals. An examination of these values will show that, for some substances, the E index is greater than the O, for others less. Now, for the first class, it follows that the velocity of the E disturbance in a plane perpendicular to the O.A. must be *less* than that of the O disturbance. This means that, for these crystals, the O wavelet is entirely outside the E. Such crystals are called *positive*. For the other class, *negative*, the converse is true, that is, the E wave-surface is entirely outside the O. Thus, a, Fig. 171, corresponds to a positive crystal, b, Fig. 171, to a negative. The student will observe that Iceland spar is a negative crystal, quartz a positive.

Table IX

Substance	n_E	n_O
Iceland spar . . .	1.486	1.658
tourmaline . . .	1.62	1.64
nitrate of soda . .	1.337	1.585
quartz.	1.553	1.544
ice	1.307	1.306
sulphate of potash	1.502	1.493

121. The Nicol Prism. — For reasons which will appear in the next chapter, it is frequently desirable to get rid of one of the doubly refracting beams. This can readily be done by the use

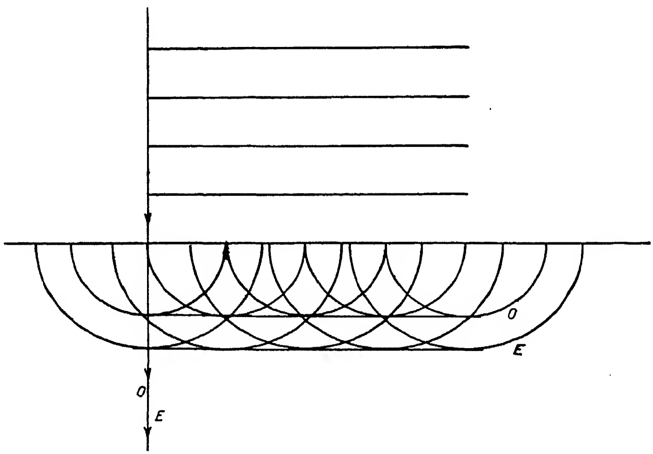


FIG. 177.

of a Nicol prism, a device in which the O beam is removed by total reflection. Its construction should be clear by a consideration of Figs. 179 and 180. To begin with, a crystal of Iceland spar, about three times as long as it is broad, is taken, and small pieces

are cut off each end somewhat as indicated in Fig. 179. In this figure $BCDF$ represents the section of the crystal marked with the same letters as in Fig. 169. After end pieces represented by FBB'

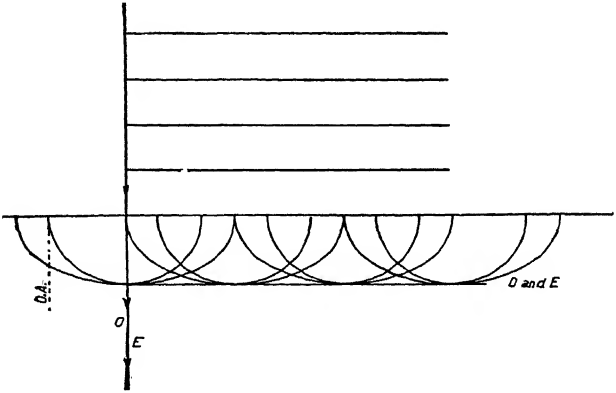


FIG. 178.

and $D'CD$ are removed, the angle $B'CD'$ is about 68° instead of the original 71° 5′ of the angle BCD. The crystal is next cut into two pieces M and N along a plane represented by $B'D'$, that is, by a

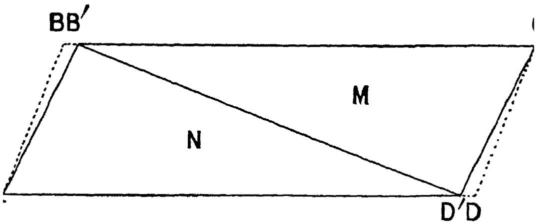


FIG. 179.

cut which passes through B' and D' (see again Fig. 169) and is at right angles to the section $B'CD'F$. These two pieces are subsequently cemented together by Canada Balsam, after which, to all appearances, the crystal is as transparent as before the cut.

There is a very important difference, however, which arises from the fact that the index for Canada Balsam, 1.55, is intermediate in value between n_E and n_O for Iceland Spar. This means that Canada Balsam is optically always less dense than the prism as far as the *O* beam is concerned, and optically more dense for the *E* beam (provided the latter does not travel in a direction too near that of the optic axis). It follows, therefore, that when the *E* beam is incident on the layer of cement, it is transmitted, since total reflection does not occur, whereas the *O* beam is totally reflected if the angle of incidence is greater than the critical angle. With the arrangement we have just described, this is ordinarily the

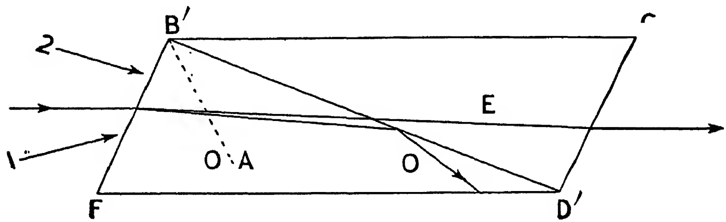


FIG. 180.

case, and consequently, as shown in Fig. 180, the *O* beam is reflected to one side and the *E* beam transmitted.

The student should note, however, that at one edge of the field of view it is *possible* to have the *O* beam transmitted, and at the other edge, to have the *E* beam totally reflected. The transmission of the *O* beam occurs if light is incident on the prism along a direction somewhat as indicated by the arrow marked 1, in Fig. 180, so that inside the prism the *O* beam strikes $B'D'$ near B' at an angle of incidence less than the critical angle. To understand why a beam, incident somewhat as indicated by the arrow marked 2, can give rise to an *E* beam which is totally reflected, the student must recall that n_E, *the* index of refraction refers only to the light travelling in a direction *perpendicular* to the optic axis. But if an *E* beam travels in a direction *parallel* to the optic axis, its velocity is the same as that of the *O* beam (see again Fig. 171). Consequently the *E* beam has a whole range of velocities between these

two extremes and, therefore, corresponding values of indices ranging from that usually called n_E to a value equal to n_O. When the E beam, therefore, travels more nearly parallel than perpendicular to the optic axis, its index may be less than that of the balsam and so even the E beam may be totally reflected. For light incident as in 2, then, we may have total reflection of the E beam incident on $B'D'$ (near D').

When approximately parallel light is transmitted through the Nicol (and this is usually the case) such edge effects have no practical importance.

122. Principal Plane. — A plane perpendicular to the face of a crystal on which light is incident and containing the optic axis, is called a *principal plane*. If one looks at the end of a Nicol

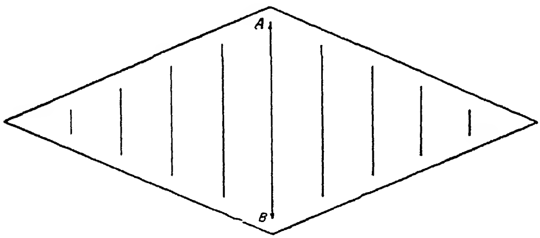


FIG. 181.

prism, a principal plane may conveniently be represented in a sectional diagram such as Fig. 181, by the line AB, joining the corners of the two obtuse angles. Since the optic axis is a direction, it must be realized that there are an infinite number of principal planes.

122A. Foucault Prism. — In this type of prism, the O beam is removed by total reflection, as in the Nicol, but the thin layer separating the two pieces in which the original crystal is cut, is air, not Canada Balsam.

Since air is optically less dense than the crystal for both the O

and the E beams, obviously total reflection is possible for both. But the critical angles are not the same, having, in fact, the values $42° 16'$ for the E beam and $37° 6'$ for the O, if we use the indices for Iceland spar given in Table X. If, then, the prism is so designed that the angles of incidence of the light striking the film are intermediate between these values, it follows that the E beam will be transmitted and the O beam totally reflected. This condition is ensured by using a crystal in which the ratio of length to width is considerably less than in the Nicol prism, thus giving an advantage which is largely commercial.

A disadvantage to the Foucault prism, which is not in general use like the Nicol, is the fact that, although the E beam is transmitted, even for it a fair amount of intensity is lost by reflection.

122B. Polaroid, a Substitute for a Nicol Prism. — Certain doubly refracting crystals exhibit to a marked degree *dichroism*, that is, a property measured by the difference in the absorption of the O and the E beams. For example, a piece of tourmaline, 1 millimetre or so in thickness, transmits only the E beam because of the complete absorption of the O beam. As a practical means of getting rid of one of the doubly refracting beams, however, tourmaline is of little use because the absorption of the E beam is a marked function of the wave-length and the emergent light is strongly coloured.

Since 1932 an excellent substitute for a Nicol prism has been available in the substance which bears the trade name *Polaroid.* Essentially this is an extremely thin layer of a crystalline substance called *herapathite* whose dichroism is so pronounced that a thickness of only $\frac{1}{200}$ inch absorbs one of the doubly refracted beams, but transmits the other to a marked degree for all wave-lengths in the visible spectrum. Herapathite is a compound of iodine and quinine sulphate which was given this name because it was discovered by W. B. Herapath as long ago as 1851. Although its dichroic properties had been known for many years, it was not until 1932 that E. H. Land invented the process which is the basis of Polaroid. Land succeeded in making a thin sheet

of a matrix, nitrocellulose, in which ultra-microscopic crystals of herapathite were imbedded *with their optic axes all parallel.* His general method is utilized by the Polaroid Corporation of Boston for the manufacture on a commercial scale of this substitute for Nicol prisms. A thin layer of Polaroid is placed between protecting plates or films, and pieces with surface areas many times greater than those of the largest Nicols may readily be obtained at a comparatively low cost.

Before giving further information about Polaroid it is necessary to explain in detail the peculiar property of the light which emerges from either a Nicol prism or a piece of Polaroid. This is the subject of the next chapter, where, it should be noted, Nicol prisms may be replaced by Polaroid.

CHAPTER XIII

PLANE POLARIZED LIGHT

123. The Meaning of Polarization of Light. — If one views an ordinary source of light through a Nicol prism, nothing peculiar is observed, no matter what the position of the Nicol. If, however, the light is viewed through two Nicol prisms, the one nearest the light being stationary, while the other is rotated, marked changes in the intensity are observed. Thus, when the principal planes of the two Nicols are at right angles, no light at all is transmitted through the second. As it is rotated, light is transmitted with greater and greater intensity until a maximum is reached when the principal planes are parallel, that is, after a rotation of 90° from the first position. As rotation is continued, the light again becomes dimmer, a second time falling to zero intensity after another 90°. Exactly the same phenomenon may be observed by leaving the Nicol next the eye fixed and rotating the other one.

One may conclude, therefore, that light which has emerged from a Nicol prism differs from ordinary light. Moreover, since the intensity of the emergent light depends on the position of the principal plane of the rotating Nicol with respect to any fixed plane of reference in space, the difference, whatever it is, must be due to a property of the beam of light which depends on its orientation with respect to such a plane of reference. What, then, can such a property be?

Before answering that question, it is well to consider a medium such as the string of a violin which can vibrate either longitudinally or transversely. If such a string is stroked longitudinally, the vibrations of all particles along its length bear the same relation to a vertical or to a horizontal, or to any other plane passing through the string. If, however, the string is stroked or plucked sideways and thus made to vibrate transversely, then the relation

of the vibration of any particle to a plane of reference through the string depends very decidedly on what that plane is. For example, if the string is vibrating in a plane perpendicular to the chosen plane of reference, no component of the vibration of the particle takes place in that plane, whereas if the plane of reference is parallel to the vibrating plane, the whole of the vibration of each particle takes place in it. If, therefore, transverse vibrations are confined to a definite plane, the component of the vibration of any particle in any arbitrary plane of reference depends on the angle which this plane makes with the vibration plane. With longitudinal vibrations, however, no such relation exists.

It is concluded, then, from the results obtained by looking through two Nicol prisms, (1) that light waves must be transverse; (2) that as a result of the passage of light through a Nicol prism, the vibrations are all in one definite plane. Light for which condition (2) is the case, is said to be plane polarized. In ordinary light, on the other hand, while the vibrations are transverse, the direction in which each particle is vibrating, is continually and rapidly changing. *On the average*, then, the magnitude of the component of a vibration in one plane is the same as for any other. In the case of a Nicol prism, polarization results because the crystalline structure forces, as it were, the vibrations all to take place in one plane. When, therefore, two Nicols are symmetrically placed (that is, with principal planes parallel), the vibrations of the light emerging from the first and incident on the second are parallel to the plane in which alone it transmits vibrations and the transmitted intensity is a maximum. Again, when the two are placed with their principal planes at right angles, there is no component of the incident light in the plane in which the second transmits vibrations, and no light emerges.

It is natural now to ask, what *is* this vibration plane, that is, how is it related to the principal plane? Anything which has been said so far is true regardless of what angle the vibration plane may make with the principal plane. Can we find what this angle is? The answer to that question can be found by a consideration of another method of obtaining plane polarized light.

124. Polarization by Scattering. — In section 114 attention was directed to the phenomenon of scattering by small particles. Suppose now that light scattered in a direction, *AB*, Fig. 182,

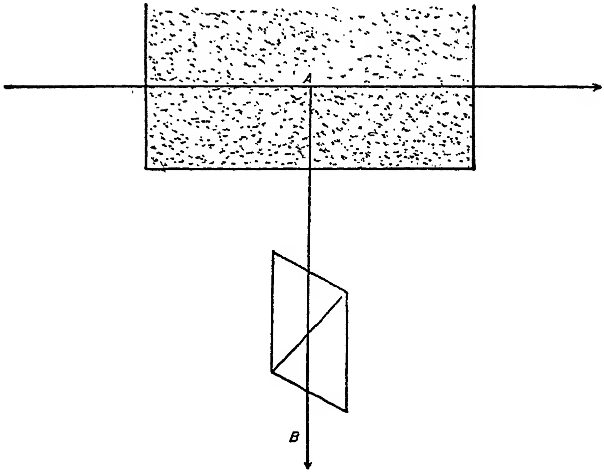


FIG. 182.

that is, along lines perpendicular to the path of the original beam, is observed *through a Nicol prism*. It is found that as the Nicol is rotated, the intensity changes from a maximum to a minimum every 90°. The scattered light is, therefore, plane polarized. Moreover, if the original beam is horizontal, and if light waves are transverse, it follows that the vibration plane for light scattered in the direction *AB* must be a vertical one. Now if the position of the Nicol prism is noted when a maximum amount of light is transmitted, it is found that the principal plane is also vertical. We conclude, therefore, that *the plane in which a Nicol prism transmits vibrations must be a principal plane.*

In the same way the vibration plane of the light transmitted by a piece of Polaroid may be located. (Alternately, the Polaroid may be used in conjunction with a Nicol prism.)

A variation of the above experiment leads to the same conclu-

sion. Suppose the *original* beam is plane polarized as a result of passage through a Nicol prism placed as N, Fig. 183. We now find, if we look in the direction AB ***with the unaided eye*** that as the Nicol is rotated, the intensity of the scattered light changes from a maximum to a minimum every 90°. Moreover, when the intensity is a maximum, the principal plane of the Nicol is vertical, while

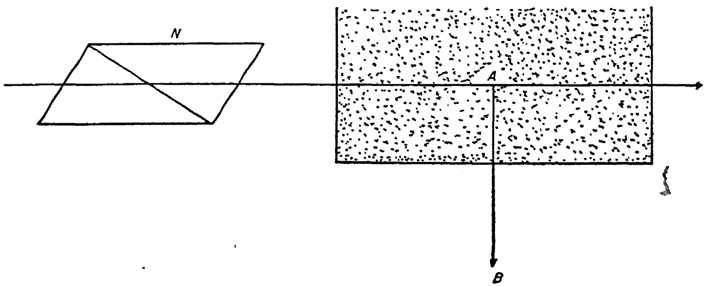


FIG. 183.

when it is a minimum, the principal plane is horizontal. The explanation of these results is at once forthcoming, if the vibration plane of light transmitted by a Nicol is a principal plane. For when the principal plane is vertical, the plane polarized beam which encounters the scattering particles causes vibrations at all places in a vertical direction, that is, along lines at right angles to the line of sight (a, Fig. 184), and such vibrations are transmitted. On the other hand, when the principal plane is horizontal, the plane polarized beam gives rise to vibrations at all places in a direction parallel to AB, and therefore, with no component perpendicular to the line of sight. (b,Fig. 184.) Because of the transverse nature of light vibrations, therefore, no disturbance is sent in the direction AB. If that were the case, we should have longitudinal vibrations giving rise to light. Once more we conclude that the vibration plane of light transmitted by a Nicol prism is a principal plane.

Before leaving the subject of polarization by scattering, it is not amiss to point out that it follows from the above considerations

that the blue light from the sky should be polarized. That this is the case can readily be verified by the simple use of a Nicol.

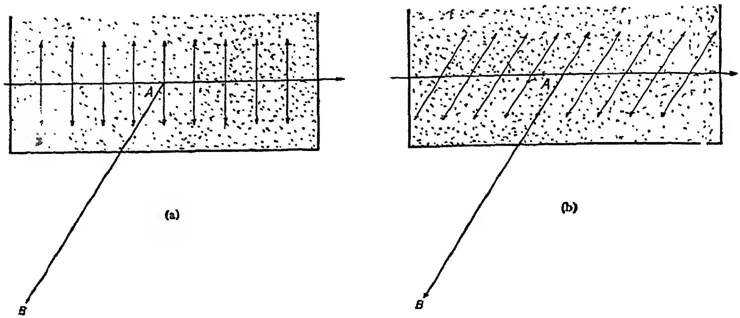


FIG. 184.

125. *O* and *E* Beams each Polarized. — It should now be abundantly clear to the student that, if we wish to test a beam of light for polarization, all that is necessary is to examine the light by a rotating Nicol. For example, if the *O* and *E* beams emerging from a piece of Iceland spar are examined in this way, it is found not only that each beam is plane polarized but that the vibration planes are at right angles. For, when the intensity of the *E* beam is a maximum, the *O* beam is absent, and on rotating the Nicol through 90° from this position, the *E* beam disappears, the *O* beam being of maximum intensity. Now, since in a Nicol prism it is the *E* beam which is transmitted, it follows that the vibration plane for *O* light is perpendicular to the principal plane, while for the *E* beam, it is in a principal plane. In the general mathematical treatment dealing with the propagation of plane waves through a crystalline medium, it is shown that, in a given direction, only waves with vibrations either in, or perpendicular to, a principal plane can be transmitted.

126. Double-Image Prism. — Suppose a beam of light is incident normally on the face *MA*, Fig. 185, of a rectangular block made of two wedge-shaped pieces of a doubly refracting

medium such as quartz. Suppose, further, that the piece MNA has been cut so that the optic axis is parallel to the surface MA and in the plane of the paper, while, for the piece NAB, the optic axis is parallel to the surface NB but perpendicular to the plane of the paper. For the first prism, then, the optic axis is perpendicular to the refracting edge (through A), for the second, it is parallel to its refracting edge (through N).

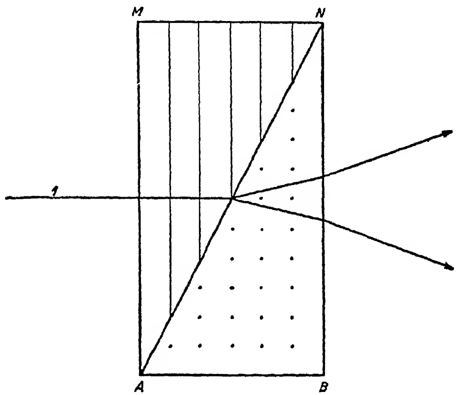


FIG. 185.

Consider, now, ray 1 incident normally on the face MA. In the prism it gives rise to O and E disturbances, which, although not separated, travel at unequal rates and with vibration planes at right angles. On entering the second prism, since the optic axis of this is at right angles to that of the first, the original E beam proceeds as an O beam, the O beam as an E beam. The velocity of one, therefore, is increased, of the other, decreased. Hence, at the oblique surface NA, the beams are refracted in opposite directions, and, on emergence from the face NB, two widely separated beams are obtained.

Such an arrangement constitutes a Wollaston double-image prism. A prism of this sort is sometimes attached to an eye-piece when one is examining phenomena which exhibit polari-

zation. For example, when the transverse normal Zeeman effect is being observed through a telescope equipped with a (properly oriented) double-image eye-piece, one of the two images seen shows the two outer components of the normal triplet, the other, the central component. (See also section 157.)

127. Polarization by Reflection. — A third important method of obtaining polarized light was discovered by Malus (France, 1775–1812) early in the nineteenth century. The method may be demonstrated by allowing a beam of ordinary light to be reflected from a pile of thin glass plates, as in Fig. 186, at an

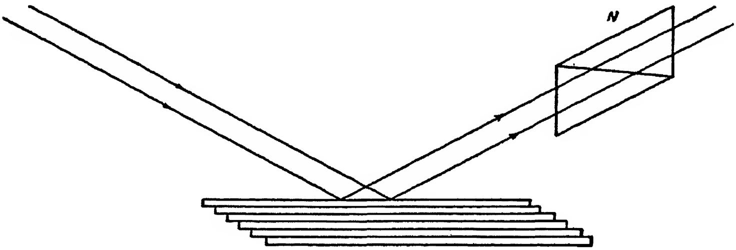


FIG. 186.

angle a little less than 60°. By examining the reflected beam with a rotating Nicol prism, it is easy to show that this light is plane polarized, with the vibration plane perpendicular to the plane of incidence. If, at the same time, the beam transmitted by the plates is tested, it too, will be found to be polarized, but with the vibration plane in the plane of incidence. Instead of glass, other transparent media may be used.

By varying the angle of incidence, in the case of a beam incident on glass, it is found that there is one angle for which polarization is practically complete. For other angles, while the rotating Nicol shows changes in intensity, at minima the light is not completely cut off. In such cases the reflected beam consists of a mixture of plane polarized and ordinary light, or is said to be partially polarized. For substances other than glass, while there is always an angle of maximum polarization, even at this angle polarization

is not always complete. Sir David Brewster (Great Britain, 1781–1868), who made measurements on the angle of maximum polarization, showed that, for any substance, its value could be found from the relation

$$\tan i = n, \tag{13.01}$$

where i is the required angle of incidence and n is the index of refraction.

If a single plate of glass is used, the reflected beam is of feeble intensity and a large portion of the transmitted light is unpolarized. When two plates are used, a second reflected beam is added to the first, thus making the reflected beam more intense, while at the same time the percentage of polarized light in the transmitted beam is increased. By using a number of plates, a strong reflected beam of polarized light is obtained, and with a sufficient number, the transmitted light becomes completely polarized. It is as if the plates sorted out the vibrations into two lots, one perpendicular to, the other in, the plane of incidence.

Pile of Plates as Analyzer. — In the case of a Nicol prism we have seen that it may be used either to produce plane polarized light, thus constituting a *polarizer*, or, as a means of examining a beam for the presence of polarization, in which case it is called an *analyzer*. The same is equally true of a pile of glass plates. It may be used either to produce plane polarized light or to analyze it. Thus, if the polarized light leaving a pile of plates P_1, Fig. 187, is allowed to fall on a second similar pile P_2 at the polarizing angle, a rotation of P_2 about an axis such that the beam is always incident at the same angle, shows marked changes in the intensity of the reflected beam B. When the planes of incidence of the two piles are parallel, the intensity of this beam is a maximum, whereas when the planes of incidence are at right angles, the intensity is a minimum. Thus, like a Nicol, a rotating pile of plates shows the presence of polarized light in a beam incident on it. A combination of a polarizer and an analyzer, such as two Nicol prisms, or a pile of plates and one Nicol, or two piles of plates, constitutes the essential parts of a *polariscope*.

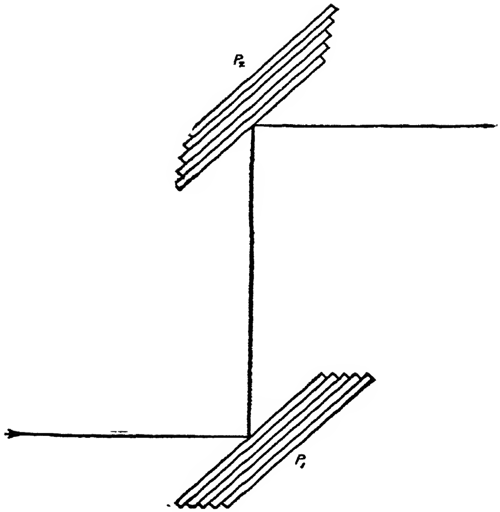


FIG. 187.

128. Plane of Polarization. — What is called the *plane of polarization* was originally defined, with reference to polarization by reflection, as that particular plane of incidence in which polarized light is most completely reflected. Now, the *vibration* plane having been shown, by means of a Nicol, to be perpendicular to the plane of incidence when light is polarized by reflection, and polarized light being most completely reflected from P_2, in the arrangement of Fig. 187, when the planes of incidence are parallel, it follows that *the vibration plane must be perpendicular to the plane of polarization*. The student must be careful to avoid identifying the two and is advised to use the term vibration plane wherever possible.

129. Law of Malus. — When a beam of plane polarized light is analyzed by either a Nicol prism or a pile of glass plates, we have seen that the intensity changes from a maximum to a minimum when the analyzer is rotated through 90°. This intensity change

must take place subject to some law. Such a law, first suggested by Malus, after whom it is called, and amply confirmed by experiment, is given by the relation

$$I_\theta = I_0 \cos^2 \theta, \qquad (13.02)$$

where I_θ is the intensity of the beam when the analyzer has been rotated through an angle θ from the position of maximum intensity (I_0). As a simple example, it is readily seen that, if a Nicol is turned 30° from the position of maximum intensity, the intensity has dropped to $\left(\frac{\sqrt{3}}{2}\right)^2$ or $\frac{3}{4}$ of the maximum.

Since at any given place in the path of a beam of light, the

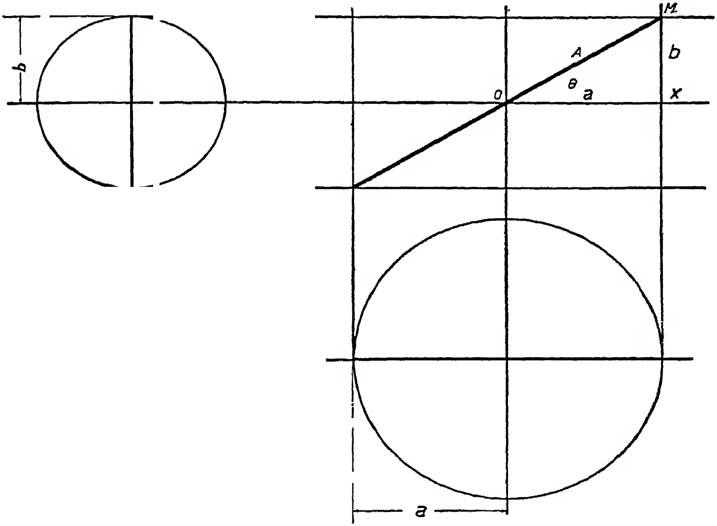


FIG. 188.

intensity is proportional to the square of the amplitude, the Law of Malus may be written

$$A_\theta = A_0 \cos \theta, \qquad (13.03)$$

where A_θ is a measure of the amplitude when an analyzer has

been turned through an angle θ from the position of maximum intensity, and where A_0 is a measure of the maximum amplitude. Now it is frequently more convenient to work with amplitudes than with intensities because of the convenient way in which amplitudes may be either combined or resolved. In section 16, for example, it was shown how readily a resultant amplitude could be found when a particle is acted on by two S.H.M. of the same period, along lines at right angles. In particular, when the phase difference is zero, the solution illustrated by Fig. 188 (and Fig. 23) shows that the resultant in such a case is a linear S.H.M. of amplitude OM, in a direction making an angle θ with the direction of the disturbance of amplitude OX along the x axis. Now,

$$\tan \theta = \frac{MX}{OX} = \frac{b}{a},$$

where b and a are the amplitudes of the individual y and x motions.

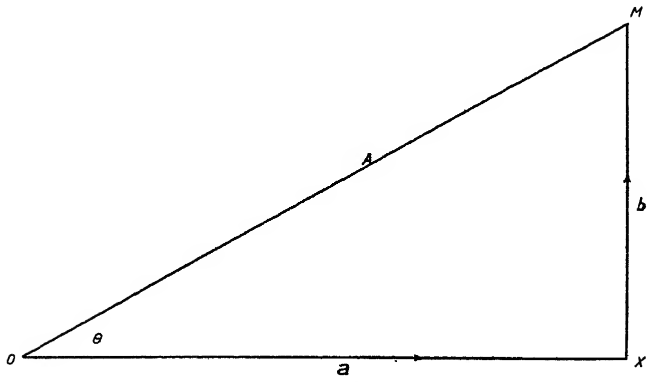


FIG. 189.

We see then, that *in the special case* of two S.H.M. at right angles, of equal periods, with zero phase difference, the resultant amplitude A is represented by the hypotenuse of a right-angled triangle, the other two sides being a and b. A may quickly be found, therefore, by the simple diagram shown in Fig. 189. The

converse is equally true, that is, *a linear S.H.M. of amplitude A may be resolved into, or is dynamically equivalent to, two linear S.H.M. along lines at right angles, one of amplitude* $a = A \cos \theta$, *and other of amplitude* $b = A \sin \theta$. In dealing with simple quantitative questions having to do with plane polarized light, this idea is extremely useful, as the solution of the following question will show.

Problem. *A beam of plane polarized light falls on a piece of Iceland spar thick enough to separate the O and the E components. If the vibration plane of the incident light makes an angle of* 20° *with the principal plane of the crystal, find the ratio of the intensity of the E beam to that of the O beam.*

If the intensity of the original beam is proportional to a^2 or $= K^2a^2$, where K^2 is a constant, then the amplitude of the original beam $= Ka$.

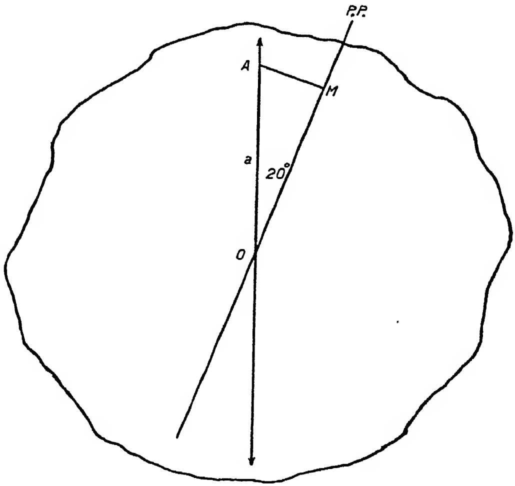


FIG. 190.

Since the crystal transmits vibrations only in the principal plane (the E beam), or perpendicular to it (the O beam), it is

natural to resolve the incident disturbance of amplitude $a = OA$, Fig. 190, into one component, $OM = Ka \cos 20°$ in the principal plane, and the other, of amplitude $AM = Ka \sin 20°$ perpendicular to the principal plane. The first component will then be transmitted as the E beam, the second as the O beam, so that we have

$$\frac{\text{amplitude of } E \text{ beam}}{\text{amplitude of } O \text{ beam}} = \frac{Ka \cos 20°}{Ka \sin 20°} = 2.75$$

and, therefore,

$$\frac{\text{intensity of } E \text{ beam}}{\text{intensity of } O \text{ beam}} \quad (2.75)^2 = 7.5.$$

129A. Advantages and Disadvantages of Polaroid. — The chief advantage of Polaroid lies in the fact that it provides us with a means of obtaining wide beams of plane polarized light at comparatively small cost. For that reason Land's invention has given popular emphasis to certain practical uses of polarizing devices. One or two of these we note.

(1) A certain portion of the light incident on many semi-smooth surfaces, such as blackboards, cover glasses of paintings, polished tables, pavements, and glazed pottery gives rise to unwanted and sometimes annoying reflected beams which are polarized to a considerable extent. For either visual or photographic purposes, these reflections may be removed by viewing the surface through a piece of Polaroid placed so that its vibration plane is perpendicular to that of the reflected light. In some cases it is convenient to use Polaroid spectacles.

(2) The problem of glare in motor-car night driving can b solved provided all cars are equipped with Polaroid over the hea lights and a Polaroid plate through which the driver views t on-coming traffic. If the vibration plane of the viewing plate perpendicular to that of the plates over the head-lights of an a proaching car, these lights are almost invisible and the cause glare is removed. At the same time objects illuminated by th driver's car are clearly seen. When the road surface is fairly horizontal, a suitable arrangement is to place all the pieces of

Polaroid on any one car with vibration planes at an angle of 45° with the vertical. On a high-crowned road, where a car near the ditch may be off horizontal, the vibration plane of the light from such a car would not be at right angles to the plane of the light transmitted through the viewing plate of an approaching car in the middle of the road, and this arrangement would not be entirely satisfactory. The difficulty arising in such a situation can be overcome by the use of circularly polarized light. This is explained in section 135A.

(3) Stereoscopic moving pictures are possible with the aid of Polaroid. Two views of the same scene are made simultaneously, one corresponding to the view seen by the right eye, the other by the left. The two pictures are projected on a screen, one through one piece of Polaroid, the second through a piece with vibration plane at right angles to the first. If an observer views the superimposed pictures through Polaroid spectacles, with the vibration plane of the piece over the right eye parallel to the vibration plane of the corresponding projected beam, and with the left spectacle similarly placed with respect to the other beam, then the effect of binocular vision is obtained and the scene appears to be in three dimensions. The material of the screen used must be of such a nature that the light is not de-polarized by reflection.

Although in most experiments, a piece of Polaroid can be used instead of a Nicol prism, it is a good substitute only over a limited range of the spectrum. This may readily be shown by the following simple experiment. Take two Polaroid plates and set them in the "crossed" position, that is, with their vibration planes at right angles, and then view through them a tungsten filament lamp of clear glass. It will be found that, although the general background is dark, the incandescent filament is seen in a purple colour. This is proof that there is transmission of light at the extreme red end of the spectrum. Actually, careful tests made with a pair of crossed plates and with monochromatic light show that the amount of light transmitted by the pair for any wavelength in the visible part of the spectrum never exceeds 5 per cent. of the incident light, and in that part of the spectrum to

which the eye is most sensitive the percentage is considerably less. At the extreme red end, however, the transmission begins to increase rapidly. In the near infra-red region the amount transmitted through crossed plates or films is just about the same as when the two pieces are parallel. A pair of crossed plates of Polaroid, therefore, may be used as an infra-red filter, that is, a device which eliminates the visible but transmits the infra-red. To test this the writer photographed on infra-red plates an incandescent lamp with a pair of crossed Polaroid plates in front of the camera. The resulting picture showed the lamp in full detail. In this connection it is worth while noting that crossed Nicol prisms give almost complete extinction over the entire spectrum as far as 200,000 angstroms or 2μ.

CHAPTER XIV

INTERFERENCE OF POLARIZED LIGHT

130. We have seen that when a beam of light is incident on a doubly refracting crystal it is in general broken into two plane polarized beams, with vibration planes at right angles, which travel at unequal rates in any specified direction. In the case of a very thin crystal, or of one cut so that light is travelling perpendicular to the optic axis (see again Figs. 176 and 177), there is no separation into two distinct beams. Because of the unequal velocities, however, on emergence from the crystal, there is a difference in phase between the two disturbances. The value of this difference in phase, for any given kind of light, depends on two things, (1) the ratio of the velocities, (2) the length of the path through the crystal. If, therefore, a thin piece of variable thickness is used, the difference in phase is different for each part of the surface from which the light is emerging. Interference phenomena may then be expected. Do such phenomena exist? Before answering this question, an examination of the following quantitative example is advisable.

Problem. *A beam of plane polarized light, for which* $\lambda = 6 \times 10^{-5}$ *cm., falls normally on a thin piece of quartz cut so that the optic axis lies in the surface. If* $n_E = 1.553$ *and* $n_O = 1.544$, *find for what thickness of the crystal the difference in phase between the E and the O beams, on emergence, is* π *radians.*

Let t = the required thickness.

Then, as far as O light is concerned, we have t cm. of crystal = $1.544\,t$ cm. of air. As far as the E beam is concerned, it is to be noted that the light is travelling perpendicular to the optic axis, and therefore the velocity is given by the value of the extraordinary index (see again section 119). Hence,

$$t \text{ cm. of crystal} = 1.553\,t \text{ cm. of air, for } E \text{ beam.}$$

Therefore, the path *difference* for the two beams

$$= (1.553 - 1.544)\ t \text{ cm. of air, that is } (n_E - n_O)\ t.$$

But the phase difference

$$= \frac{2\pi}{\lambda} \cdot \text{path difference}$$

$$\therefore \qquad = \frac{2\pi}{\lambda} \cdot (1.553 - 1.544)t$$

The phase difference, however $= \pi$

$$\therefore \qquad \frac{2\pi}{\lambda}(1.553 - 1.544)\ t = \pi$$

or $$t = \frac{6 \times 10^{-5}}{2 \times .009} = 0.033 \text{ mm.}$$

A second solution.

Since λ_O, the O wave-length in the crystal $= \dfrac{6 \times 10^{-5}}{1.544}$,

and λ_E, the E " " " " " $= \dfrac{6 \times 10^{-5}}{1.553}$,

$\therefore$ the number of O waves in t cm. $= \dfrac{1.544\,t}{6 \times 10^{-5}}$

and " " " E " " t cm. $= \dfrac{1.553\,t}{6 \times 10^{-5}}$,

or the difference in the number of E and of O waves

$$= \frac{t}{6 \times 10^{-5}}(1.553 - 1.544).$$

But, since the phase difference is π, the difference in the number of O and of E waves must be $\frac{1}{2}$ or any odd multiple of $\frac{1}{2}$.

$\therefore \quad \dfrac{t}{6 \times 10^{-5}}(1.553 - 1.544) = K \cdot \frac{1}{2}$, where K is any odd integral member.

Hence, just as before,

$$t = 0.033 \text{ mm. when } K = 1.$$

131. Supposing now that we had a crystal of this thickness, would the resultant of the two beams, differing in phase by π, give rise to darkness? That this cannot be the case should be at once evident if the student recalls that the vibrations of the two disturbances are at right angles. For two S.H.M. acting on a single particle along lines at right angles could never bring that particle to rest, no matter what the phase difference. As the work of section 16 shows, in such cases we have resultant paths which may be along straight lines, or in ellipses, or occasionally in circles — never zero disturbance. We see, then, that we cannot expect to have interference phenomena with such an arrangement.

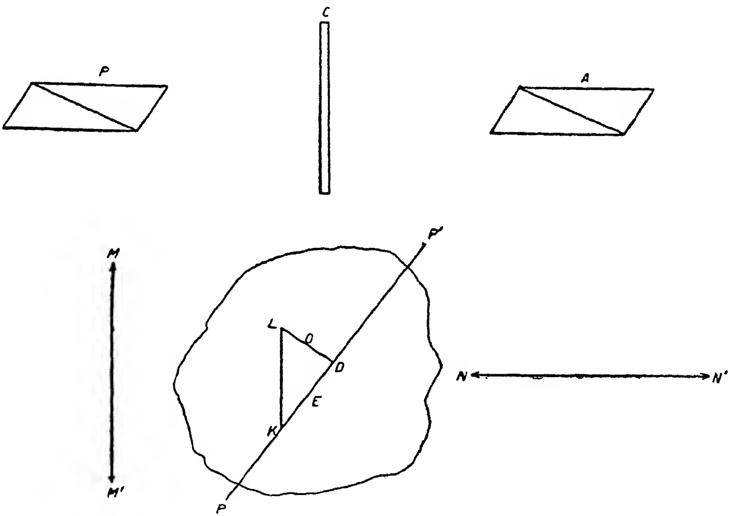


FIG. 191.

If, however, the light emerging from the crystal passes through a Nicol prism (or any analyzer), then a portion of each disturbance is transmitted with vibrations in the principal plane of the Nicol, and we then have the necessary conditions for interference. The exact arrangement necessary for the observation of interference effects with polarized light, as well as for the explanation of what

is observed, will perhaps be clear with the aid of Fig. 191 and Fig. 192. In Fig. 191, P and A represent a polarizer and an analyzer between which a thin piece of doubly refracting crystal C, is placed; the line MM' represents the vibration plane of the plane polarized light leaving P, NN', that of the light transmitted by A. In the diagram these two vibration planes are at right angles, in which case the analyzer and the polarizer are said to be in the *crossed* position. In the absence of the crystal, obviously no light at all is transmitted when the analyzer and polarizer are crossed.

The diagram immediately below C represents the face of the crystal whose principal plane passes through PP'. If, then,

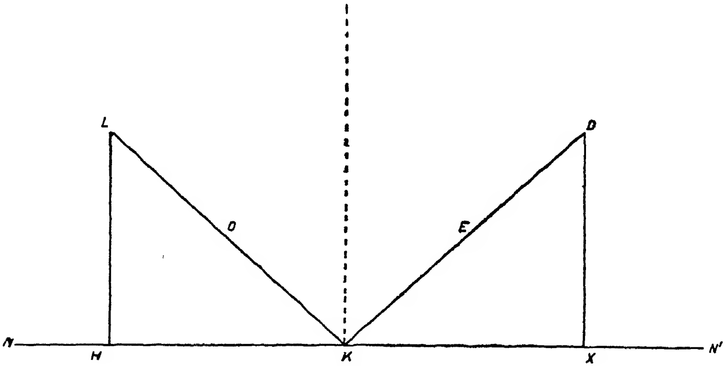


FIG. 192.

LK represents the amplitude of the incident light, it will be broken into an E component of amplitude KD with vibrations in the principal plane, and an O component of amplitude LD with vibrations perpendicular to the principal plane. On emerging from the thin crystal there will be a difference in phase between the two components, whose magnitude, once again, for a given wavelength, and a given direction, depends upon the thickness of the crystal. On striking the analyzer, a portion of each will be transmitted with amplitudes as represented in Fig. 192. Thus, since the E disturbance of amplitude KD is equivalent to one of amplitude KX with vibrations parallel to NN' together with another of

amplitude XD with vibrations perpendicular to NN', the first of these only will be transmitted (with amplitude KX). Similarly, since the O disturbance of amplitude LK is equivalent to one of amplitude HK, parallel to NN', and LH perpendicular to NN', HK alone will be transmitted by the analyzer. We have then, finally, two disturbances, one of amplitude HK, the other of amplitude KX, emerging from the analyzer with a phase difference depending on the thickness of the crystal. Interference is now possible.

If the light is monochromatic and the crystal is of variable thickness, we see on a screen on which the emergent light falls, some places bright, others dark. For example, if a piece of quartz in the form of a wedge is used, and it is so placed that the O and the E beams are of equal intensity (that is, so that PP' makes an angle of 45° with NN', in Fig. 191), there are alternately bright and dark regions parallel to the thin end of the wedge. An actual photograph of such fringes is reproduced in Fig. 5, Plate IV.

If white light is used, the pattern on the screen is generally brilliantly colored. The reason for this is the same as that given on page 203, when the question of interference colours observed with soap films was under discussion. Destructive interference (that is, zero intensity) in general can be obtained for only one wave-length at a time, for, if the phase difference is π for λ_1, it will not be so for most wave-lengths differing from λ_1. In the light pattern, therefore, corresponding to a particular thickness, while there may be a place of zero intensity for λ_1, (and possibly other wave-lengths), for most wave-lengths the intensity will not be a minimum. This place on the screen will accordingly be coloured, for it will represent white light minus a certain percentage of the constituent wave-lengths.

Complementary Colours. — An interesting variation of the above arrangement is obtained when a piece of ordinary Iceland spar (or a double-image prism) is substituted for the analyzing Nicol. If the piece is so chosen that the O and E beams, while separated, overlap to some extent, in the overlapping portion there must always be white light, for it represents the *total* energy

of the beam incident on the analyzer. Accordingly, if the thin crystal is chosen of uniform thickness so that a uniform patch of coloured light is obtained with an ordinary analyzer, the change to the piece giving two emergent beams will show complementary colours, with white in the overlapping portion.

To obtain such interference phenomena the analyzer and polarizer need not be necessarily in the crossed position. Indeed, if the analyzer is rotated, while the colour pattern changes, it is always present, except for four critical positions in each revolution. That there should be such critical positions it is easy to see with reference to Fig. 191. For when the analyzer is being rotated and the vibration plane NN' becomes parallel to KD, the E beam is transmitted with undiminished intensity, while the O beam of amplitude LD is completely cut off. Only one beam being then transmitted, interference is impossible. Similarly, when NN' becomes parallel to LD, the E beam is completely cut off, the O alone being transmitted, and again interference is out of the question.

If the analyzer and polarizer are left in fixed positions, and the crystal is rotated about an axis parallel to the beam of light, the same result is obtained, although for a slightly different reason. Thus, on rotating the crystal, when PP' becomes parallel to MM' the whole of the beam of amplitude LK is transmitted as an E beam, the O beam being absent. Again, when PP' is perpendicular to MM', there is no component of LK in the principal plane, hence the E beam is absent, the disturbance being transmitted with amplitude LK as an O beam. In both these cases, there being only one beam leaving the thin crystal, interference is impossible.

131A. Interference Phenomena with Scattering Medium as Analyzer. — An interesting variation of the above arrangement may be had by using as analyzer a tank of water with a suspension of scattering particles (see section 114). In the following experiments, in which thin pieces of doubly refracting material of uniform thickness are needed, cellophane is satisfactory.

(1) A horizontal, approximately parallel, beam of unpolarized light falls on an opaque screen containing two or three windows of Polaroid mounted with vibration planes horizontal. Immediately beyond the Polaroid is a second opaque screen with corresponding windows of different thicknesses of cellophane. For the best results the cellophane should be cut so that the principal plane makes an angle of 45° with the vertical. The resulting beams traversing the tank placed in the path of the light are beautifully coloured when viewed from a direction at right angles to the beam. The particular shade of each beam of course depends on the thickness of the corresponding window of cellophane.

(2) If the polarizing screen is replaced by a single fair-sized piece of Polaroid, such as the standard 4-inch square, and if this is rotated through 90°, the colours of the beams change to their complements. During the rotation a critical position for which each beam loses its colour can be observed.

(3) If a single window of cellophane is used, complementary colours may be observed simultaneously by mounting a plane mirror above the tank. If the mirror is suitably inclined the observer can view the beam both horizontally, by direct light, and vertically, by light reflected from the mirror. The two beams thus seen are always complementary in colour.

(4) Adjacent beams of complementary colours may be obtained by using as polarizer two pieces of Polaroid placed one above the other, with their line of separation horizontal, but with the vibration plane of one vertical, the other horizontal. When some of the light emerging from each piece traverses a cellophane window (preferably made of several layers), the colours in the upper and lower halves of the beam in the tank are complementary.

132. The whole question may conveniently be discussed analytically and somewhat more generally. In Fig. 193, OP represents the vibration plane of light transmitted by the polarizer, while OP'' represents the principal plane of the crystal, with OR a plane at right angles to the latter.

Let $POP'' = \alpha$; $AOP'' = \beta$, and let $OK = a$, represent the

amplitude of the light incident on the thin crystal. OK is then the equivalent of LK of Fig. 191.

We have then, leaving the crystal, an E beam of amplitude a

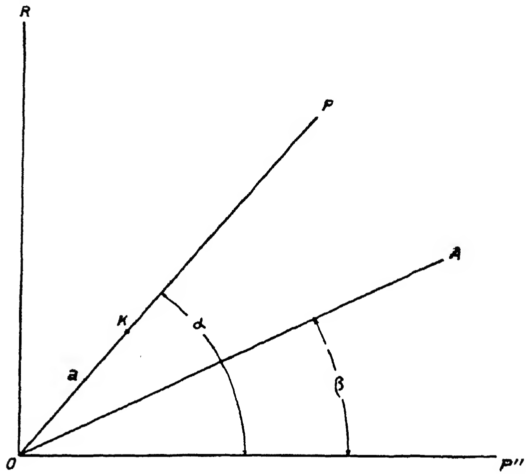


FIG. 193.

$\cos \alpha$, with vibrations parallel to OP'', and an O beam of amplitude $a \sin \alpha$, with vibrations perpendicular to OP'' or parallel to OR.

When these two components strike the analyzer, they give rise to

$a \cos \alpha \cos \beta$, vibrations parallel to OA,
$a \cos \alpha \sin \beta$, vibrations perpendicular to OA,
$a \sin \alpha \sin \beta$, vibrations parallel to OA,
$a \sin \alpha \cos \beta$, vibrations perpendicular to OA.

Thus we have *transmitted* by the analyzer two components of amplitudes $a \cos \alpha \cos \beta$ and $a \sin \alpha \sin \beta$, between which there is the difference in phase resulting from the unequal velocities of the E and the O beams in the thin crystal. If this phase difference $= \delta$, the final resultant amplitude A of the light transmitted is then found just as has been indicated on page 185. Thus

$$A^2 = a^2 \cos^2 \alpha \cos^2 \beta + a^2 \sin^2 \alpha \sin^2 \beta + 2\, a^2 \cos \alpha \cos \beta \sin \alpha \sin \beta \cos \delta$$

$$= a^2 \Big[\cos^2 \alpha \cos^2 \beta + \sin^2 \alpha \sin^2 \beta + 2 \cos \alpha \cos \beta \sin \alpha \sin \beta - 4 \cos \alpha \cos \beta \sin \alpha \sin \beta \sin^2 \frac{\delta}{2} \Big]$$

$$= a^2 \Big[(\cos \alpha \cos \beta + \sin \alpha \sin \beta)^2 - 4 \cos \alpha \cos \beta \sin \alpha \sin \beta \sin^2 \frac{\delta}{2} \Big]$$

$$= a^2 \Big[\cos^2 (\alpha - \beta) - \sin 2\alpha \sin 2\beta \sin^2 \frac{\delta}{2} \Big]. \qquad (14.01)$$

Since the intensity is proportional to the square of the amplitude, this expression enables us to find how the intensity varies due to changes in δ, as well as in α and β. We note the following cases:

(1) When the analyzer and polarizer are in fixed positions, that is, when α and β remain constant.

The intensity then varies with the value of $\sin^2 \frac{\delta}{2}$ and δ, as we have already seen, varies with the thickness of the crystal and the wave-length. The general results already discussed at once follow.

(2) When the analyzer and polarizer are crossed.

In this case $\alpha - \beta = 90°$, and we have

$$A^2 = a^2 \Big[0 - \sin 2\alpha \sin (2\alpha - 180°) \sin^2 \frac{\delta}{2} \Big]$$

$$= a^2 \sin^2 2\alpha \sin^2 \frac{\delta}{2}. \qquad (14.02)$$

(3) When the analyzer and polarizer are parallel.

In this case $\alpha = \beta$, and we have

$$A^2 = a^2 \Big[1 - \sin^2 2\alpha \sin^2 \frac{\delta}{2} \Big]. \qquad (14.03)$$

It will be noticed that the sum of the two expressions (14.02)

and (14.03) is always a^2, regardless of the value of δ or of α. This is just what should be expected, for the sum of these two beams must represent the total intensity, provided there is no actual absorption by the analyzer. In the case where an ordinary doubly refracting crystal replaces the analyzer, and complementary colours are obtained, the overlapping portion represents the above sum, and, as we have seen, is white.

133. Test of Strain in Glass. — Ordinary glass is isotropic and exhibits no sign of double refraction. When glass is subject to great strain, however, it becomes doubly refracting. A convenient test of the presence of such strains is found, therefore, by placing the specimen between a crossed analyzer and polarizer. If the glass has strains, it is doubly refracting, two beams are transmitted, and the strains are rendered evident by the appearance of interference colours.

An important application of this fact has been made to determine the distribution of stresses in solid opaque materials used in engineering construction. To do so, models of the material to be examined are made in glass or in a transparent plastic substance like celluloid or bakelite. The models are then placed between analyzer and polarizer, and from a study of the colour pattern corresponding to the strains in the transparent material, the actual stresses can be determined.

The examination of stresses in this way has developed into a branch of engineering called *photoelasticity*. Its advantages are self-evident, especially in problems where stress analysis by theoretical considerations is difficult or impossible.

134. Elliptically and Circularly Polarized Light. — We return now to a consideration of the nature of the light emerging from a thin crystal such as we have been discussing. Since we are dealing with cases where the O and the E beams are not separated, although vibrating in planes at right angles, we have to do with an application of the work discussed when considering the composition of two S.H.M. at right angles.

Diagrams such as those shown in Figs. 23, 24, 25, 26 and 27 tell us that the resultant in such cases may be a linear motion, or a motion in an elliptical, or in a circular path. The first of these results when the phase difference between the two components is any multiple of π. To obtain circular vibrations, two conditions must be fulfilled: (1) the phase difference must be $\frac{\pi}{2}$; (2) the amplitudes of the individual beams must be equal. Such conditions obtain, (1) if a piece of thin crystal is chosen of such a thickness t that the required phase difference results, that is, such that

$$\frac{2\pi}{\lambda} \cdot (n_E - n_O)t = \frac{\pi}{2}, \qquad \text{[see section 130]}$$

or

$$t = \frac{\lambda}{4(n_E - n_O)},$$

and (2), if this crystal is so placed in the path of the plane polarized beam that its principal plane makes an angle of 45° with the vibration plane of the incident light.

A crystal of such a thickness is called a *quarter-wave plate.*

Elliptical vibrations represent the general case when the two beams with vibrations at right angles have any phase difference other than a multiple of π — except the special case giving rise to circular vibrations.

135. We see then that in the beam of light leaving a thin crystal on which plane polarized light is incident, we may have resultant vibrations confined to a definite plane, or taking place in elliptical paths or in circular. In other words, we may have either plane polarized light, or elliptically polarized light, or circularly polarized. The question then arises, can one analyze a beam of light and detect the presence not only of plane polarization but also of the other kinds? The answer to that we shall briefly discuss, considering the following four cases.

(1) As we have seen several times, if light is completely plane polarized, a rotating Nicol will show regions of maximum intensity alternating with minima *where complete extinction occurs.*

(2) If the beam consists of a mixture of plane polarized and ordinary light, the rotating Nicol again shows places of maxima and minima intensities, but at the latter the intensity does not drop to zero. Only the polarized part of the beam can be completely cut off.

(3) Suppose a beam of light is completely circularly polarized. As the converse of the work on the composition of S.H.M. at right

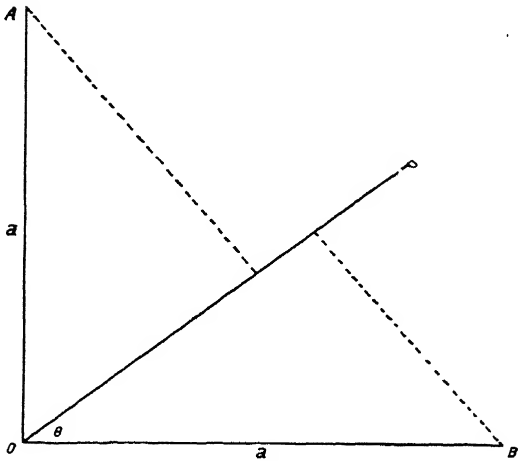


FIG. 194.

angles shows, *a circular vibration may be resolved into or is dynamically equivalent to two linear vibrations of equal amplitudes along any two lines at right angles, with a phase difference equal to* $\frac{\pi}{2}$. When, therefore, a beam of circularly polarized light falls on a Nicol prism, there is no change in the intensity of the emergent light as the Nicol is rotated. That this is so, should be clear from Fig. 194, where $OB = OA = a$ represent the amplitudes of the component linear vibrations, as well as their vibration planes, while OP is the vibration plane of the light transmitted by the Nicol. If the angle $POB = \theta$, we see that the Nicol transmits

two components, one of amplitude $a \cos \theta$, the other, of amplitude $a \sin \theta$. Since the phase difference between these two is $\frac{\pi}{2}$, the resultant amplitude A is therefore given by (see p. 39 and p. 185)

$$A^2 = a^2 \cos^2 \theta + a^2 \sin^2 \theta + 2a^2 \cos \theta \sin \theta \cos \frac{\pi}{2} = a^2.$$

Hence, regardless of the value of θ, the emergent intensity is always the same. The rotating Nicol, therefore, shows no change in intensity, and hence does not distinguish between ordinary and circularly polarized light.

Suppose, however, that we introduce into the path of the circularly polarized beam a quarter-wave plate. Then a further phase difference of $\frac{\pi}{2}$ is introduced between the components so that on emergence from this plate, the phase difference is either 0 or π. In either case, plane polarized light results, and now, when the beam falls on a rotating Nicol, complete extinction occurs for certain positions. If, however, a quarter-wave plate is introduced into the path of ordinary light, plane polarized light does not emerge and the rotating Nicol shows no changes in intensity.

(4) Consider next a beam of elliptically polarized light. Although an elliptical vibration, as we have seen, is the general resultant of any two linear motions at right angles, it can always be built up by combining two linear along lines parallel to the major and to the minor axis, and differing in phase by $\frac{\pi}{2}$ (see Fig. 26). Conversely, an elliptical motion is dynamically the equivalent of two linear vibrations of unequal amplitude, along lines parallel to the axes, and differing in phase by $\frac{\pi}{2}$. When, therefore, elliptically polarized light falls on a rotating Nicol, it is not difficult to see that there are changes in intensity, a maximum existing when the principal plane is parallel to the major axis of the ellipse, a minimum when it is parallel to the minor axis. But this is just

what is observed when a beam consists of a mixture of plane polarized and ordinary light. To distinguish between the two, a quarter-wave plate may be used in the same way and for the same reason as in the analysis of circularly polarized light.

More exact and more detailed analysis of the general nature of beams of light with respect to the kind and degree of polarization can be made by the use of a Babinet Compensator. Essentially this consists of two small-angled wedge-shaped pieces of quartz, placed as in Fig. 195, so as to form a rectangular piece of uniform thickness. The pieces are cut so that in each the optic axis is parallel to the surface, but in *A*, in the plane of the paper, in *B*, perpendicular to this plane. If a ray of polarized light is incident normally in the direction of the arrow, in general an *E* and an *O* beam result, between which a phase difference is introduced

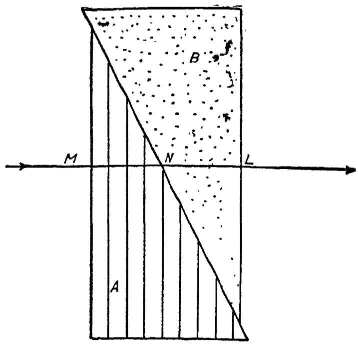


FIG. 195.

$$= \frac{2\pi}{\lambda}(n_E - n_O)\, MN, \text{ in the first piece,}$$

and
$$= \frac{2\pi}{\lambda}(n_E - n_O)\, NL, \text{ in the second piece.}$$

Now since the principal planes of the two pieces are at right angles, that component which traverses *A* as an *E* beam (that is, with vibrations in the principal plane) becomes the *O* beam for the second piece. The velocities, therefore, are interchanged in the second piece, and so the resultant phase difference is given by the difference of the above expressions, that is, by

$$\frac{2\pi}{\lambda}(n_E - n_O)(MN - NL).$$

It follows that if the piece B can slide with respect to the other, the path length NL can either be increased or decreased, and so the phase difference on emergence at L can be made any desired value. Consequently the light emerging at L can be made either linearly or elliptically or circularly polarized.

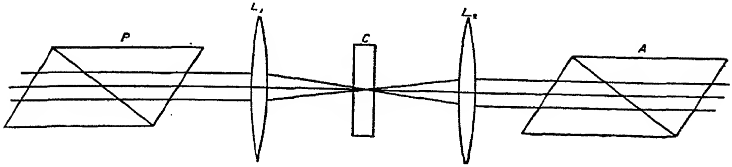


FIG. 196.

If circular polarized light is wanted, or if it is desired to introduce at L a phase difference exactly equal to $\frac{\pi}{2}$, this device has an advantage over a quarter-wave plate in that the latter gives the required phase difference exactly for only one wave-length, whereas the compensator can be set accurately for any wave-length.

As the use of this device in the general analysis of the state of polarization of a beam of light involves principles similar to those discussed in connection with the quarter-wave plate, further details will not be given.

135A. It was pointed out in (2) of section 129A that, when two approaching cars are not on the same horizontal level, the arrangement there described is not entirely satisfactory. The difficulty may be overcome by the use of circularly polarized light. Suppose that the pieces of Polaroid over the head-lights and the Polaroid viewing screen are covered with a thin plastic layer whose thickness is so chosen that it is a quarter-wave plate for a wave-length in the region of maximum sensitivity of the eye. The beams leaving the head-lights are then circularly polarized provided the quarter-wave plate is mounted so that condition (2), section 134, is fulfilled. Moreover, circularly polarized light emerges whether

the car is on absolutely level ground or not. When this circularly polarized light falls on the viewing piece of Polaroid similarly covered (on its outer surface) with a quarter-wave plate, no light is transmitted to the eye of the observer in an approaching car. As explained at the end of (3) in section 135, the second quarter-wave plate converts the circularly polarized beam into plane polarized light and this is cut off by the Polaroid plate next the eye.

136. Interference with Highly Convergent Light. — In the arrangement for the study of interference with polarized light

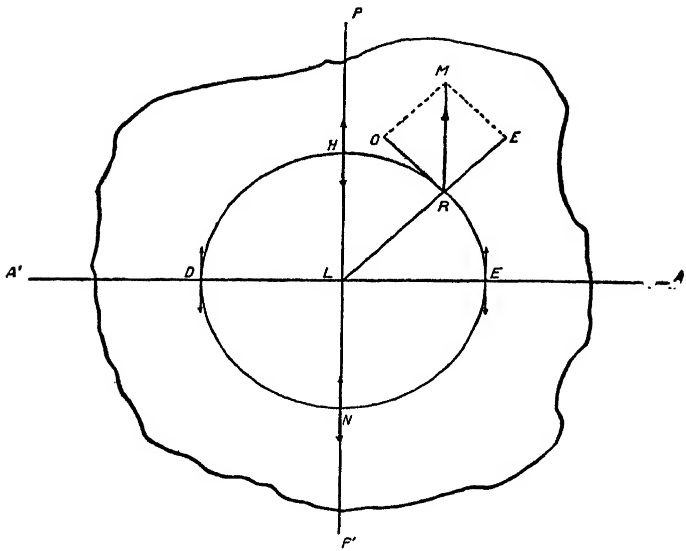


FIG. 197.

which we have been considering, approximately parallel rays have been used. We shall now briefly consider the type of interference phenomenon when the incident light is highly convergent. Suppose, then, by means of an arrangement such as is shown in Fig. 196, where L_1 and L_2 represent lens *systems*,

a uniaxial crystal C is placed in the path of a beam of convergent white light, plane polarized by the Nicol prism P, and subsequently analyzed by the analyzing Nicol A. Consider first the special case when the analyzer and polarizer are crossed so that in Fig. 197, which represents the face of the crystal, the vibration plane of the incident light is represented by PP', that of the light transmitted by the analyzer by AA', at right angles to PP'. In this figure L represents the point at which the axis of the incident beam intersects the face of the crystal.

If now the crystal is cut with the optic axis perpendicular to the face on which the light is incident, along the axis of the beam of light (through L) the O and the E beams travel with equal velocity and no phase difference results. Indeed, in this direction, for a reason which will appear presently, there is only one beam transmitted through the crystal. Consider, however, any point R where the incident light is not normal to the face of the crystal. The original disturbance of amplitude RM, with vibrations parallel to PP', is at this point broken into two, an E beam with vibrations in the principal plane and therefore parallel to RE (since the principal plane contains the optic axis and the incident ray), and an O beam with vibrations perpendicular to RE, or along RO. The direction of a ray incident at R not being parallel to the optic axis, the two disturbances travel through the crystal with unequal velocities, and on emergence a phase difference results. Since the velocity of the E ray varies with the direction and is the same for all lines making equal angles with the optic axis, the magnitude of the phase difference for a crystal of uniform thickness is the same for all points on a circle with centre L and radius LR. When, therefore, the emergent light is analyzed by the analyzing Nicol, the colour resulting from interference is the same for all points on this circle, except for the four points H, D, N and E, which are always black. At the points H and N, the vibration plane of the incident light is in a principal plane and so the O beam is entirely absent. When, therefore, the light from such points falls on the analyzer with its vibration plane perpendicular to PP', none at all is transmitted. Since the same

can be said with regard to any point on the line PP', we see that the complete light pattern transmitted will be crossed with a black line or band parallel to PP' Similarly, at the points D and E and all points on the line AA', the vibration plane of the incident light is perpendicular to the principal plane, and so is transmitted unchanged as an O beam. Corresponding to these points, therefore, a black band parallel to AA' is observed in the transmitted light.

Again, since for all other points on the crystal face, the phase difference between the O and the E disturbances steadily changes as circles of radii greater or less than LR are taken, in addition to the dark bands, there is a series of circular coloured interference fringes. A photograph of the appearance is given in Fig. 3, Plate VI.

If the analyzer is placed so that its vibration plane is parallel to that of the polarizer, the single beam for all points along PP' and AA' is completely transmitted, and the black bands are replaced by bright ones.

If the crystal is cut with the optic axis parallel to the surface, general analysis shows that the fringes are now hyperbolic, while for an oblique section they are elliptic or hyperbolic.

136A. In a recent note in *Nature* Buckley, an English crystallographer, has drawn attention to a simple arrangement whereby, with the aid of Polaroid plates or films of fairly wide aperture, interference patterns with highly convergent light may be obtained. In this arrangement light from a projection lantern falls on the first piece of Polaroid, the polarizer, and is then strongly converged by a lens system such as the Abbe microscope condenser. The doubly refracting crystal is placed as in Fig. 196, but unlike the arrangement in that illustration the only additional apparatus necessary is the analyzer, in this case the second piece of Polaroid, placed in the path of the diverging beam, and a screen. A good-sized image of the interference pattern is seen on the screen.

PLATE VI

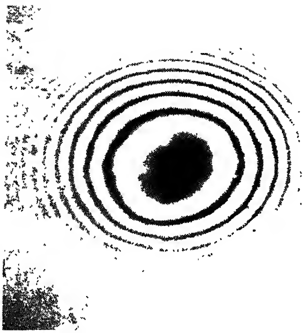


FIG. 1. Newton's Rings.

FIG. 2. Stark effect. (Foster)

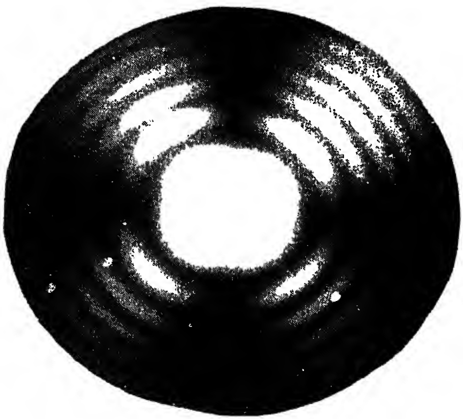


FIG. 3. Interference pattern for highly convergent polarized light and quartz cut perpendicular to the optic axis.

CHAPTER XV

ROTATORY POLARIZATION

137. Rotation of the Vibration Plane. — Suppose, with a monochromatic source such as a sodium flame, a polarizer P, Fig. 198, and an analyzer A have been placed *in the crossed position.* If, now, a piece of quartz Q, a millimetre or two thick, and *cut with its face perpendicular to the optic axis*, is inserted between the polarizer and the analyzer, it is found that light is transmitted by the latter. On the other hand, rotation of the analyzer through a definite angle θ (about 22° if the quartz is 1 mm. thick and yellow light is used), once more completely cuts off the light. It follows, therefore, that the vibration plane of the light leaving the quartz

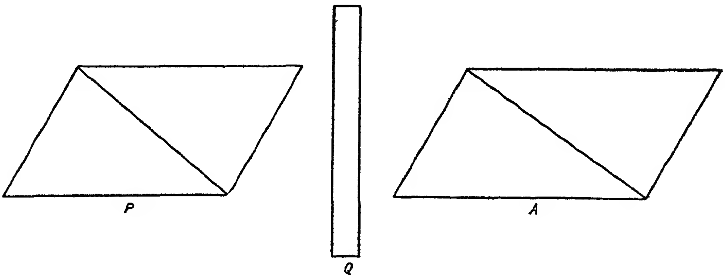


FIG. 198.

must make an angle equal to θ with that of the light incident on it. In other words, as a result of the passage of the light through the crystal along the optic axis, the vibration plane has been rotated through a definite angle. If yellow light is replaced by red, the same general result is obtained except that the angle of rotation is less, about 17° for red light of wave-length 6500 Å, for a crystal 1 mm. thick.

Such an experiment, first performed by Arago (France, 1786–1853) in 1811 illustrates the phenomenon of rotatory polarization.

An extended examination of substances having the property of rotating the plane of polarization has shown that it is possessed by, (1) crystalline quartz along the optic axis; (2) certain liquids such as turpentine, as well as solutions of sugar, tartaric acid; (3) certain isotropic substances when placed in a strong magnetic field.

Quartz and optically active substances in general, that is, liquids, and those which in solution have this property, may rotate the vibration plane either to the right or to the left, and consequently are classified as dextro-rotatory or laevo-rotatory. The same substance may exhibit both kinds, some crystals of quartz, for example, rotating to the left, others to the right.

Concerning the actual amount of rotation, the following simple laws have been experimentally found:

(1) In the case of quartz and pure liquids, for a fixed wave-length, the angle of rotation is directly proportional to the thickness. For solutions, the angle for a given thickness is proportional to the concentration of the optically active substance.

(2) For a fixed thickness, the angle of rotation increases as the wave-length decreases, roughly inversely as the square of the wave-length.

Thus, for a crystal of quartz 1 mm. thick, we have the values for the angle of rotation given in Table X.

Table X.

Rotation of Vibration Plane by 1 mm. quartz, at 20° C.

Wave-length	6708	6563	5893	5351	4861	4102
Rotation	16.4°	17.3°	21.7°	26.5°	32.7°	47.5°

It will be noted that these values apply only at 20° C. For other temperatures, a correction must be made; thus for D light, the rotation at any temperature $t°$ may be found from the relation

$$\theta = 21.^\circ72\,[1 + 0.000147\,(t - 20)].$$

138. An interesting illustration of the variation in the rotation with wave-length is had when the monochromatic source is replaced by white light. In this case, rotation of the analyzer can never produce extinction of the light, but for each position, a coloured beam emerges. The colour results from the fact that, for every position of the analyzer, there is, in general, one wave-length for which the vibration plane is exactly at right angles to that of the light transmitted by the analyzer. This wave-length, therefore, is completely cut off, while a certain percentage of the intensity of all other wave-lengths (whose vibration planes are not perpendicular to that of the analyzer) is transmitted. The resultant light transmitted is, therefore, coloured.

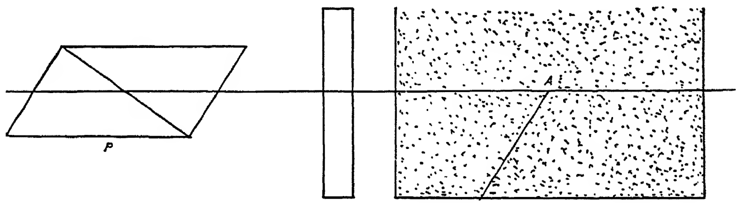


FIG. 199.

The variation of the amount of rotation with changing wave-length can be nicely shown by using a tank of water containing fine scattering particles in suspension somewhat after the manner described in section 114. Suppose that, with white light as source, we place between a polarizing Nicol P, Fig. 199, and the tank, a piece of quartz Q cut perpendicular to the optic axis. The plane polarized beam of white light incident on Q then gives rise, on emergence, to vibration planes, in different azimuths for different wave-lengths. Consequently, if an observer is looking at right angles to the beam traversing the tank, that is, along the line AB, there is in general, always some wave-length for which the vibrations are parallel to AB, and another wave-length for

which the vibrations are perpendicular to this direction. Light corresponding to the first wave-length is, therefore, entirely absent, to the second, present with maximum intensity, and the resultant beam observed is coloured. As the polarizing Nicol is rotated, the colour shade alters, because that wave-length for which no light is transmitted along AB is then continuously changing. Of course, the exact shade of the beam seen in direction AB depends also on the excess scattering in this direction of the shorter wave-lengths.

This experiment may be varied so as to give, if successfully done, a striking and very beautiful phenomenon, by removing the quartz and using in the tank, or better, in a long tube, a strong sugar solution containing scattering particles in suspension. In this case, since the vibration plane of any particular wave-length is being continuously rotated as the light traverses the solution, the resultant shade seen when looking at right angles is also constantly changing along the length of the tube. The net result is to give a coloured spiral appearance to the whole tube.

139. Tint of Passage. — Another illustration of colour phenomenon due to the same general cause is obtained by taking a piece of quartz of such a thickness that the rotation for a wave-length in the greenish-yellow region, that is, in the brightest part of the spectrum, is 90°. If, for example, the thickness is 3.75 mm., the rotation per millimetre is

$$\frac{90}{3 \cdot 75} = 24^\circ,$$

and Table X shows that the corresponding wave-length is about 5600 A. Suppose, then, that such a piece is inserted between a polarizer and an analyzer, the vibration plane of the former being represented by POP', Fig. 200. In the same figure, then, the vibration plane for light of wave-length 5600 on emergence from the crystal is represented by MOM', the angle POM being equal to 90°. For a longer wave-length, 6000, for example, the vibration plane is represented by NON', the angle PON being con-

siderably less than 90°. For a shorter wave-length such as 5200, the vibration plane is rotated through a greater angle, and hence is represented by ZOZ'.

If, now, the analyzer is turned so that its vibration plane is parallel to that of the polarizer, light of wave-length 5600 is completely cut off, the resultant shade of the transmitted light being a greyish-violet. A slight rotation one way from this position evidently increases the transmitted intensity of 6000 and other wave-lengths towards the red end of the spectrum, but decreases that of 5000 and the shorter waves at the blue end. The greyish-violet tint is then replaced by a deep pink. A slight rotation in the opposite direction, however, decreases the intensity of the longer waves, increasing that of the shorter, and so causes a quick change from the greyish-violet to a deep blue. As the contrast between pink and blue is very marked and takes place as the result of a slight rotation, the grey-violet shade is called the *tint of passage* or *sensitive tint*. Because of this fact, even with white light as a source, it is possible to set an analyzer in a critical position, if a piece of quartz of such a thickness is placed before it.

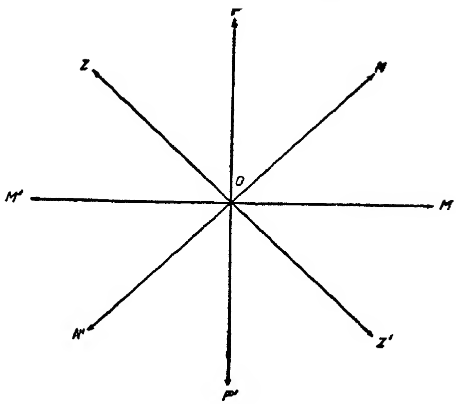


FIG. 200.

If the simple piece of quartz is replaced by two adjoining pieces of equal thickness, one dextro-, the other laevo-rotatory, a still more sensitive arrangement, called a *bi-quartz* is provided. In this case, when the polarizer and analyzer are parallel, the sensitive tint is obtained in each piece. A slight rotation of the analyzer, however, increases the intensity of the longer wave-lengths in

one, at the same time increasing that of the short wave-lengths in the other, and thus causes one piece to become pink, the other blue. Rotation in the opposite direction interchanges the colours. A slight rotation through the critical position, therefore, when the two adjoining pieces have the same shade, causes such a very marked change in the two halves that a setting in this position can be accurately made. Reference will again be made to this point in the next section.

140. Analysis of Optically Active Substances. — Since the angle of rotation of an optically active substance in solution is proportional to the concentration, a quantitative means is thus provided for studying the strength of such solutions. In the case of sugars, this fact is particularly important because "commercial sugars are bought and sold and the import duties collected upon tests" of the amount of rotation. "There are upwards of one hundred sugars known and nearly all of those in solution rotate the azimuth of a rectangular vibration in one direction or the other." (Skinner.) For the exact measurement of the amount of rotation, the observer uses a *saccharimeter*, another name given a polarimeter when used for measuring optical rotations of this kind. In the simplest form, all that is required is a polarizer, an analyzer whose exact position can be read from a circular scale, and, between the two, a tube with transparent ends containing the optically active solution. In this form, with monochromatic light as source, the analyzer is set for extinction before and after the introduction of the solution, the angle of rotation of course being equal to the difference in the two readings. The position of extinction, however, cannot be set exactly enough for accurate measurements, and, in actual practice, settings are generally made by what is called the brightness half-shade method. In essence, this consists in so altering the simple polarimeter arrangement that a field of view is obtained with at least two contiguous parts, the final setting of the analyzer being made when these are of equal brightness.

A particular example, in which a Cornu-Jellett half-shade

Nicol is used to obtain the necessary two portions, should make the principle clear. This form of Nicol is made by taking an ordinary Nicol prism, and cutting along its full length a thin *wedge-shaped* piece, somewhat as represented by $ABCDEF$ in Fig. 201, after which the two parts 1 and 2 are cemented together. As a result the principal planes of these two parts are now slightly

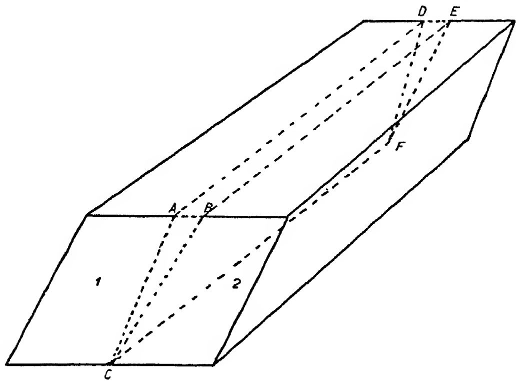


FIG. 201.

inclined to one another, and therefore, also, the vibration planes of the light transmitted by each. If, then, a beam of plane polarized light is incident on a Cornu-Jellett Nicol used as an analyzer, and it is set so that no light is transmitted by part 2, a feeble intensity will be transmitted by part 1. Thus, if in Fig. 202, XOX' represents the vibration plane of the incident light, and, for a certain position of the analyzer, OM_1 represents the vibration plane of light transmitted by 1, ON_1 that by 2, part 2 is dark (since the angle N_1OX is 90°), while some light is transmitted by 1. Again, if the analyzer is rotated through the angle M_1ON_1, so that ON_1 (or OM_2) represents the vibration plane of 1, and ON_2 that of 2, part 1 is now dark, part 2 transmitting a little light. In between these positions there is obviously one place for which each part transmits a feeble intensity of the same brightness, and because of the contrast between the two halves resulting

from a slight rotation either way, this position can be set with great accuracy. With monochromatic light and this form of Nicol as analyzer, accurate measurements of angles of rotation can be made.

From the work of the preceding section, it should be clear that with white light as a source and the use of a bi-quartz between an

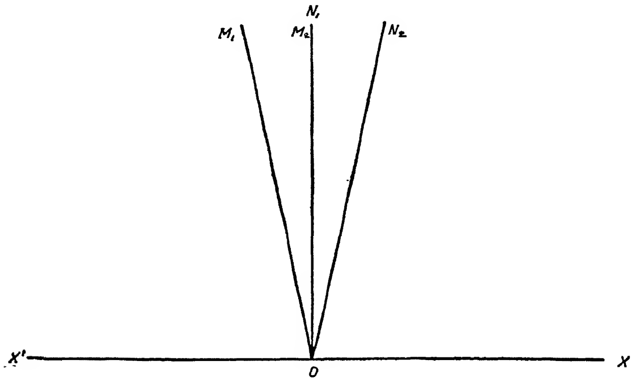


FIG. 202.

ordinary analyzer and polarizer, settings of the analyzer may be made. In this case, measurements of the amount of rotation are evidently made for that particular wave-length whose vibration plane is rotated through 90°. What this wave-length is can be found by examining the light transmitted by either half of the bi-quartz with a spectroscope when the principal planes of analyzer and polarizer are parallel.

141. Specific Rotation. — To compare the optical activity of different substances, a unit must be adopted. This is done by defining the term *specific rotation* in the following way:

$$\text{Specific rotation} = \frac{\text{angle of rotation for a thickness of 10 cm.}}{\text{density of solution in grams per cc.}}$$

A few actual values, for a temperature of 20° C., are given in Table XI. (Taken from Kaye and Laby's Tables).

Table XI.

Specific Rotations of Certain Substances.

Substance	solvent	specific rotation	condition
Glucose	water	−94.4° after 7 mins. −51.4° after 7 hrs.	when 4 gms. dissolved in 100 gms. of solution
Turpentine	pure liquid	−37°	
Nicotine	pure	−162°	at temperature 10° to 30° C.
Rochelle salt	water	+ 29.73° − .0078 c.	c = concentration in gms. per 100 cc. of solution.

The plus sign indicates dextro-rotatory, that is, anti-clockwise when looking along the beam away from the source, the minus sign, laevo-rotatory.

142. Rotation by Magnetic Field. — In 1845, Michael Faraday (England, 1791–1867) showed that when a piece of glass of high refractive index traversed by a beam of plane polarized light, is placed in a strong magnetic field, a rotation of the vibration plane occurs. Since the original discovery, further experiments by Faraday and by later investigators have shown that a similar phenomenon is observed, although sometimes in a very slight degree, by many solids, liquids and gases. The actual angle of rotation θ is directly proportional to, (1) the strength H of the magnetic field, (2) the length l, of the path traversed in the field. The amount of rotation may, therefore, be expressed by the law

$$\theta = cHl,$$

where c is a constant, called Verdet's constant.

If the direction of the light is reversed, the sense of rotation is also reversed. Thus if a clockwise rotation occurs when light travels in the direction of the lines of force, the rotation is contra-clockwise when the beam is reversed. In consequence of this, if a beam is reflected back and forth along the lines of force, the amount of rotation is greater, the greater the number of times the path is traversed. For this reason rotations may be measured for substances which exhibit this property to only a very slight extent. In the case of optical active substances, the rotation is always in the same sense regardless of the direction of the beam of light, in consequence of which a reversal of the beam annuls the original rotation.

143. Explanation of Rotation of Vibration Plane. — We have seen in Chapter XIV, that for thin crystals an incident beam of plane polarized light can emerge as a plane polarized beam with the vibration plane turned through an angle. For this to be the case, the phase difference between the O and the E components vibrating at right angles must be some odd multiple of π, that is, it occurs only for crystals of very special thicknesses. In rotatory polarization, however, plane polarized light emerges for any thickness of the substance, with an angle of rotation which steadily increases with increasing thickness. We must, therefore, look elsewhere for the explanation.

There is another way in which plane polarized light can result when two disturbances act on a particle, and that is, when the two are each circularly polarized but in opposite senses. Thus a rectilinear vibration along a line may be the resultant of two circular, and, conversely, a rectilinear vibration may be resolved into or is dynamically the equivalent of, two equal and opposite circular vibrations. This being the case, it is possible to explain the phenomena we are discussing on the supposition that we have to do with a class of substances which may be said to be *circularly doubly refracting* in that they transmit only circular vibrations, and these with velocities which are unequal for the two senses of rotation. This was the explanation first proposed by Fresnel,

who showed in a way presently to be described, that in quartz along the optic axis, there are two such disturbances.

First of all it is necessary to see that, according to this explanation, the difference between dextro- and laevo-rotatory substances arises from the fact that in the one, the contra-clockwise vibration is propagated the faster, in the other, the clockwise vibration. Consider then, a particle executing a rectilinear motion parallel to the line *OM*, Fig. 203. Such a motion is the equivalent of two

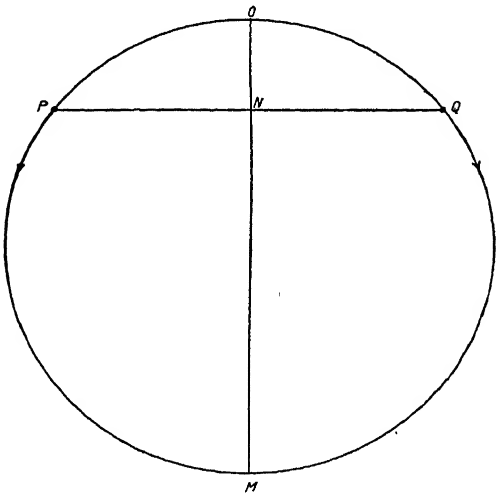


FIG. 203.

equal and opposite circular motions of the same period, for, as far as any displacement perpendicular to *OM* is concerned, the effect of one is always equal and opposite to that of the other. For example, the displacement *PN* neutralizes *QN*. Conversely, the resultant of two circular motions of the same period acting on a single particle, is a rectilinear vibration, along *OM* if the two circular disturbances pass each other (when imagined acting independently) at *O* and *M*.

Suppose, now, that plane polarized light, with vibration plane

parallel to OM, Fig. 203, is incident on a medium which transmits only circular vibrations and these with unequal velocities. The original vibration is then replaced by two equal and opposite circular motions, which, at the instant the disturbance strikes the medium may be considered to be in position O. If then the two travel through the medium at different rates, on emergence, at the instant the contra-clockwise disturbance is represented by

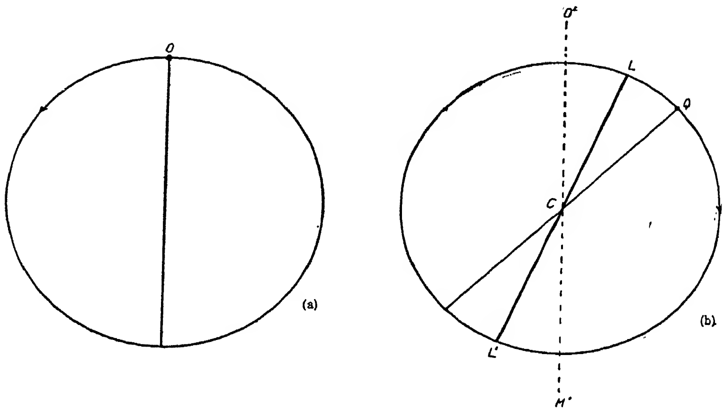


FIG. 204.

the particle at O, (a) Fig. 204, the particle for the clockwise will be at some other position, such as Q, (b) Fig. 204. The two disturbances on emerging will continue to travel at equal rates, and hence will combine into a rectilinear vibration along the line LCL', where LC bisects $O'CQ$, for at right angles to this line there can be no resultant displacement. As a result of the unequal velocities in the medium, therefore, the vibration plane has been rotated through the angle $O'CL$. Moreover, it is not difficult to see that if the velocity of *propagation* (not the period of rotation) of the clockwise motion is the slower, the rotation of the vibration plane is to the right (looking in the direction of propagation). Conversely, if the clockwise motion is the faster, the rotation of the vibration plane is in the opposite sense.

As already mentioned, Fresnel demonstrated experimentally the existence of two disturbances in quartz, travelling with unequal rates along the optic axis. To do so, he took several pieces of both dextro- and laevo-quartz, in the form of prisms, placing them together, alternately right and left, to form a rectangular solid as in Fig. 205. In each the optic axis was parallel to the sides

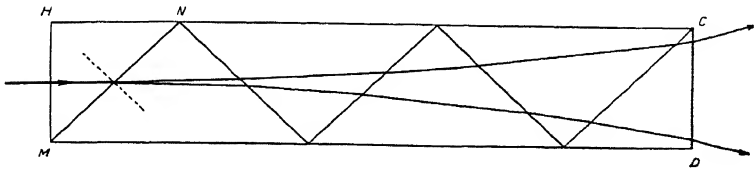


FIG. 205.

CH and *MD*. On the above theory, then, if plane polarized light is incident normally on the face at the left, two disturbances travel in this direction with unequal velocities. On striking the inclined surface *MN* and entering the second prism, the disturbance which was faster in the first prism now becomes the slower. Hence, by the ordinary laws of refraction one beam is bent away from the normal to *MN*, the other towards it. Since at the other inclined surfaces similar effects take place, it is not difficult to see that two distinctly separated beams should emerge from the right face. Fresnel showed that this actually was the case.

144. That the phenomenon of rotatory polarization is absent in fused quartz is experimental evidence that it is in some way related to the crystalline structure of this substance. In the case of liquids and of solutions of optically active substances, no crystalline structure is, of course, possible, but for these it is found that the molecules possess an asymmetric carbon atom. Moreover, as such molecules, for some substances, can occur in two forms, one of which is the mirror image of the other, it is not surprising that different specimens of the same substance can exhibit, sometimes dextro-, sometimes laevo-rotations.

PROBLEMS

1. What is Huygens' explanation of the difference between positive and negative uniaxal doubly refracting crystals?

(b) What is meant by the extraordinary index of refraction of a uniaxal crystal?

(c) Show how you would determine the vibration plane of the light leaving a Nicol prism.

2. Plane waves fall *normally* on a piece of Iceland spar cut with optic axis parallel to the surface. By making a diagram to illustrate Huygens' explanation of double refraction, prove that the O and the E wave-fronts will emerge travelling in the same direction.

3. Show, by means of a diagram, that for one particular plane, the sine law of refraction holds for both O and E beams.

4. When two Nicol prisms, with principal planes making an angle of 30° with each other, are altered so that the angle becomes 45°, how much has the intensity of the emergent light altered?

5. A narrow beam of light falls on a piece of Iceland spar that separates the emergent beams. If the light passes through a second thick crystal of the same substance (the principal planes of the two crystals making an angle of 30° with each other), and is received on a screen, describe the appearance seen, and find quantitatively the relative intensities of the images observed.

6. Two sources of light are viewed, one after the other, through two Nicol prisms (polariscope arrangement). It is found that the intensities of the emergent beams in the two cases are equal when the angles between the principal planes of the Nicols are 30° and 45° respectively. Compare the intensities of the original sources.

7. (a) A horizontal beam of light traverses a tank of water containing scattering particles and is viewed through a Nicol prism along a line at right angles to the beam. Explain for what position of the Nicol the light transmitted is a maximum.

(b) How much will the intensity decrease if the Nicol is now rotated through 30°?

8. Plane polarized light falls on a piece of thin doubly refracting crystal, the vibration plane making an angle of 30° with the principal plane of the crystal. If the difference in path of the O and the E beams on emergence is $\lambda/12$, find graphically the nature of the resulting vibration.

9. When an O and E beam of polarized light emerge from a piece of double refracting crystal, so thin that there is no separation of the beams, why is interference never obtained?

10. Plane polarized light is incident on a thin piece of quartz cut with faces parallel to the axis. What is the least thickness of the crystal for which the difference in phase of the two beams is 60°?

$n_O = 1.5442, \qquad n_E = 1.5533, \qquad \lambda = 5.89 \times 10^{-5}$ cm.

11. What conditions must be fulfilled so that when plane polarized light falls on a thin piece of doubly refracting crystal (*not* cut perpendicular to the optic axis) plane polarized light emerges, (1) with vibration plane in the same direction, (2) with vibration plane in a different direction.

12. Given $n_E = 1.5533$; $n_O = 1.5442$, what thickness of thin doubly refracting crystal would be required to make a quarter-wave plate for sodium light?

13. Describe a physical means of obtaining circular vibrations in light.

14. (a) Find the least thickness of a thin piece of doubly refracting crystal (cut with faces parallel to the optic axis) so that incident plane polarized light emerges circularly polarized?

$n_O = 1.5442 \qquad n_E = 1.5533, \qquad \lambda = 5.89 \times 10^{-5}$ cm.

(b) How must the crystal be placed in (a) to obtain circularly polarized light?

15. Plane polarized monochromatic light ($\lambda = 6 \times 10^{-5}$ cm.) of amplitude a, falls on a piece of thin doubly refracting crystal with its vibration plane making an angle of 30° with the principal plane of the crystal.

(a) What is the ratio of the intensities of the resulting O and E beams?

(b) If the optic axis lies in the surface of the crystal, what is its least thickness so that the O and E beams on emergence combine to form *plane* polarized light?

16. Plane polarized light falls normally on a plate of quartz with faces parallel to the axis. If the vibration plane of the incident light is at an angle of 30° to the principal plane, calculate the ratio of the intensities of the transmitted ordinary and extraordinary rays.

17. In the above question, if the crystal is 1 mm. thick, what is the difference of phase upon emergence of the ordinary and extraordinary rays of sodium light? Index for D line for E ray = 1.5533 and for O ray = 1.5442.

18. Explain the cause of the colour pattern observed when a thin piece of doubly refracting crystal is placed between an analyzer and a polarizer and the light allowed to fall on a screen.

19. When a very thin piece of a doubly refracting crystal is placed between two Nicol prisms, in general the emergent light is brilliantly

coloured. As the analyzing Nicol is rotated, in certain critical positions no colour is seen. Explain why this is so.

20. When a piece of quartz whose faces are cut perpendicular to the optic axis is placed between two Nicol prisms, different colours are observed as one Nicol is rotated. Explain.

21. If, for a certain wave-length, the rotation of the vibration plane of polarized light due to 1 mm. of quartz (cut perpendicular to the optic axis) is 20°, for what thickness of quartz will no light of this wave-length be transmitted when the piece is placed between two Nicol prisms with parallel principal planes? Explain your answer.

22. A parallel beam of plane polarized monochromatic light of wave-length in air = 5.890×10^{-5} cm. falls normally on a thin piece of doubly refracting crystal of thickness 0.01618 mm. cut with optic axis parallel to the surface. If $n_E = 1.5533$ and $n_O = 1.5442$, and if the vibration plane of the light incident on the crystal makes an angle of 20° with its principal plane, find the nature of the vibrations of the emergent light.

23. A beam of parallel rays traverses two Nicol prisms, with their principal planes parallel. Through what angle must one of them be rotated so that the emergent intensity is reduced to $\frac{1}{4}$ of the original emergent intensity?

24. Three Nicol prisms are placed in a row and a horizontal beam of light passed through them. What is the ratio of the intensity of the emergent light to that of the incident when the principal plane of the first is vertical, that of the second makes an angle of 10° to the left of the vertical (looking along the light), and that of the third an angle of 10° to the right of the vertical?

25. A parallel beam of plane polarized light for which $\lambda = 6 \times 10^{-5}$ cm. is incident normally on a piece of doubly refracting crystal in the form of a wedge so cut that the optic axis is parallel to the thin edge of the wedge. If the angle of the wedge is $\frac{1}{2}$° find the distance from the edge to the place along the wedge at which the emergent light is again plane polarized.

CHAPTER XVI

THE ELECTROMAGNETIC THEORY OF LIGHT

145. In the preceding chapters of this book, particularly in those dealing with interference, diffraction and polarization, much has been said about waves and vibrating particles. That the propagation of light is subject to the laws of wave-motion has been pointed out many times. Moreover, since the conception of wave-motion is invariably associated with a periodic vibration of a portion of some medium, an all-pervading ether has been postulated. The vibrations of the ultimate particles of this hypothetical medium may then be considered as constituting light disturbances. Throughout the greater part of the nineteenth century the ether was supposed to have the properties of an elastic solid, a conception which played a very important part in the development of the subject of light. Although this elastic solid theory of light ultimately was abandoned, it is necessary to examine one or two features in connection with it.

In the first place the student should recall that in an elastic solid, when a particle is displaced from its normal position, stresses or what we may call *restoring forces* are brought into play. If the displaced particle is released, and the displacement is not too large, it will eventually return to its normal position. Moreover, for all cases with which the theory of elasticity has to do, Hooke's Law gives the relation between the resulting stress and the displacement (or strain). According to this law, the stress is proportional to the strain, the factor of proportionality being called an *elastic modulus* or coefficient of elasticity. In an elastic solid, two kinds of strain are possible. (1) All the particles in a thin "slice" may be displaced with reference to a parallel layer, constituting what is called a *shear*. When a rod, fixed at one end, is twisted at the other, this is the kind of strain which results. (2) An elastic solid may be uniformly compressed or dilated. As

Hooke's Law is applicable in either case, there are two corresponding coefficients of elasticity, n, the shear or rigidity or distortion modulus, and k, the coefficient giving the ratio of the stress to a strain which is a uniform dilatation.

If a disturbance is created at a local region in an elastic solid, a wave motion will travel outwards in a manner much as we have described by the elementary method of Chapter II. An exact treatment of the problem, in which the dynamical equations of motion of an individual particle of the solid are set up, shows that there are two resulting wave disturbances, (1) one with transverse vibrations travelling with a velocity equal to $\sqrt{\frac{n}{\rho}}$; (2) a second, with longitudinal vibrations travelling with velocity $\sqrt{\frac{k + \frac{4n}{3}}{\rho}}$, where ρ is the density of the elastic solid. *Now in light the experimental facts discussed under the subject of polarization, have shown us that longitudinal vibrations cannot exist.* Here, then, is a first difficulty, if ether has the properties of an elastic solid. What becomes of the longitudinal vibration? The difficulty was removed by considering that the longitudinal vibration travelled with infinite velocity, or in other words, as will be seen from the above expression for the velocity, that the ether was incompressible.* There then remained only the transverse wave travelling with a velocity which depended on the elasticity and the density. Making use of such a conception, the great theoretical physicists of the nineteenth century, like Green, Stokes and Rayleigh, sought to deduce the known properties of light such as reflection, refraction, and double refraction. Although their work forms an important part of the theoretical development of the subject of light, the elastic solid theory led to difficulties and even to contradictions, and with the advent of Maxwell's famous Electromagnetic Theory, was discarded.

* The late Lord Kelvin showed that the difficulty could be removed and the theory developed by assuming a contractile ether in which the longitudinal velocity was zero.

146. The Electromagnetic Theory. — To understand the basic ideas in connection with this theory some knowledge of the principles of electricity is necessary. Faraday, in dealing with the attraction and repulsion between electrified bodies introduced the conception of tubes of force, or better, Faraday tubes, to visualize strains in an ether surrounding an electrified body. It was impossible for him, as for most Anglo-Saxon minds, to think of an electrified body repelling or attracting another at some distance from it, with nothing taking place in the intervening space. Electric attraction and repulsion, therefore, were considered to be the result of strains transmitted through an ether, and Faraday tubes were used, not only to give the mind a picture of such strains, but also, after Faraday's time, as a means of attacking electric problems quantitatively. Before the time of Maxwell (Great Britain, 1831–1879), this ether was not the luminiferous ether, the subjects of electricity and of light constituting two very distinct branches of physics.

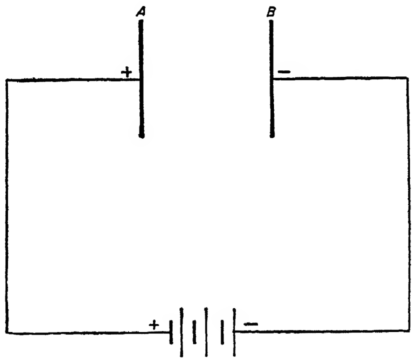


FIG. 206.

The great unification which we are about to explain was the result of the work of Maxwell, who set out to find out the properties of a medium which would transmit electric action. In establishing his mathematical equations, Maxwell introduced the somewhat revolutionary conception of *displacement* currents. According to this, when Faraday tubes are being established in free space, in air or any dielectric, there exists a momentary current which, while it lasts, has all the properties of a steady current. The idea will perhaps be clear by a consideration of Fig. 206 where A and B represent the plates of a condenser joined to the terminals of a

battery. In the static condition, the condenser plates are charged, one positive, the other negative, and Faraday tubes exist in the region between them. In this case there is, of course, no current, for we have what is ordinarily called an open circuit. But, according to the new view introduced by Maxwell, during the short initial time in which the condenser plates are being charged, or the Faraday tubes established, there exists in the dielectric between the plates a displacement current whose magnitude is proportional to the rate at which the number of tubes is increasing. Moreover, just as a steady current is surrounded by a steady magnetic field, so Maxwell assumed that the momentary displacement current gave rise to a *momentary* magnetic field in its neighborhood. If now it is recalled that an induced electromotive force results whenever there is a *change* in a magnetic field, it will be seen that the momentary magnetic field must give rise to an electric intensity. This electric intensity will then cause another momentary displacement current which, in its turn, will cause another momentary magnetic field, and so on. It follows, therefore, that granted an initial displacement current such as Maxwell postulated, an electromagnetic disturbance should be propagated from the region of the original displacement. All this Maxwell put in the form of equations, by means of which it is not difficult to show that the velocity of such a disturbance is equal to $\frac{c}{\sqrt{\mu k}}$, where c, the electromagnetic constant, is the number of electrostatic units of electricity in one electromagnetic unit quantity, where k is the dielectric constant of the medium, and μ the magnetic permeability. In free space, where $k = 1$ and $\mu = 1$, we have $v = c$.

The value of c can readily be found by standard *electrical* means. For example, the quantity of electricity stored in a condenser can be determined in electrostatic units from the product of the potential difference between its plates multiplied by the capacity calculated from its dimensions; and in electromagnetic units, by observing the current in a galvanometer when the condenser is rapidly charged and discharged an observed number

of times a second. Results obtained, whether by this or other means, give the value of $c = 3 \times 10^{10}$ cm. per second. But the observed velocity of light is 3×10^{10} cm. per second. The electromagnetic disturbance predicted by Maxwell, therefore, travels at the same speed as that of light.

Maxwell's equations also show that in such a disturbance the waves are transverse, as is the case for light. It is natural, therefore, to draw the conclusion that light is an electromagnetic phenomenon. The student, however, must guard against thinking that the electromagnetic theory in any way undermines the wave theory. It throws overboard the conception of mechanical vibrations of ether particles, but the periodicity of a light disturbance remains as before a fundamental characteristic. The ether, too, is still the seat of the disturbance, which, however, is now considered to be a periodic change in electric (and magnetic) intensity.

147. When the electromagnetic theory was developed by mathematical physicists, it was found that the laws of reflection and of refraction could readily be deduced from Maxwell's equations; that the facts of double refraction could be satisfactorily explained — in short, that the serious difficulties encountered in using an elastic solid theory disappeared when this powerful theory was put to the test. Nor should it be forgotten that two hitherto unrelated branches of science — light and electricity — were by means of it brought together by an underlying fundamental conception.

Direct experimental verification also followed, two outstanding examples being provided by the work of Boltzmann (Germany, 1844–) and Hertz (Germany, 1857–1894). According to the expression deduced by Maxwell, the velocity of an electromagnetic disturbance in a medium with dielectric constant $= k$ and permeability $= 1$ is given by the relation

$$v = \frac{}{\sqrt{k}}.$$

It follows that

$$\frac{\text{velocity in free space}}{\text{velocity in this medium}} = \sqrt{k}$$

or, if n is the index of refraction, that

$$n^2 = k. \tag{16.01}$$

Here, then, is a relation which may readily be put to an experimental test as far as light is concerned, for it is a simple matter to measure n by optical means, and k by electrical. That it is true for some media was pointed out by Boltzmann in 1873. Some actual numerical values are given in Table XII, where the indices are for yellow light.

Table XII

	n	$\sqrt{k}$
air	1.000294	1.000295
hydrogen	1.000138	1.000132
carbon dioxide	1.000449	1.000473
carbon monoxide	1.000346	1.000345

In other cases, such as water, where $n = 1.33$ and $k = 81$, there is no agreement whatever. This lack of agreement can be explained, however, when the fact of dispersion is taken into consideration — that is, the fact that different wave-lengths travel at different rates in a dielectric. Reference will be made to the theory of dispersion later, but here it may be indicated that in obtaining the value of k for water, the electrical field is static, a condition corresponding to infinite wave-length, whereas n as we have seen in Chapter VI (Cauchy's formula) is a function of the wave-length, and in light has values corresponding to short wave-lengths.

148. The Work of Hertz. — About the year 1885 remarkable experimental confirmation of the truth of the electromagnetic theory was given by Hertz, who generated electromagnetic waves by purely electrical means and showed that they possessed many of the properties of light waves. In generating electrical waves, Hertz made use of the well-known fact that the discharge of a con-

denser in a low resistance circuit is oscillatory. In such a case the periodic time is given by the relation

$$T = 2\pi \sqrt{CL},$$

where C is the capacity of the circuit, L its inductance. When a Leyden jar is discharged, the apparently single spark which we see is in reality a succession of sparks (as can be shown by a rapidly revolving mirror), corresponding to the successive alternating currents during the process of discharge. According to

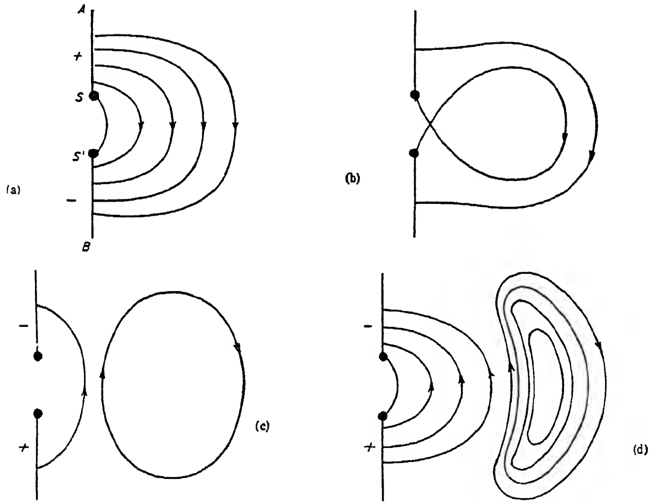


FIG. 207.

Maxwell's ideas electromagnetic waves should spread out from the region of the spark.

By the aid of Faraday tubes we may form a useful picture of the generation of such waves. Suppose, for example, that A and B, (a), Fig. 207 * are metallic plates joined to the terminals

* Figs. 207 and 208 are reproduced in part from "The Principles Underlying Radio Communication" by the kind permission of the Signal Corps, U. S. Army.

of a static electric machine, so that when the potential difference is sufficiently great a spark passes between the two small spheres S and S'. Just before the spark passes, if A is positively charged, B negatively, Faraday tubes pass from A to B, somewhat as represented in (a), Fig. 207. The arrow indicates the direction of the electric intensity, that is, the direction in which a small body carrying a unit positive charge would move if placed in the electric field. When the spark first passes and the discharge begins to take place, many of the Faraday tubes disappear, the

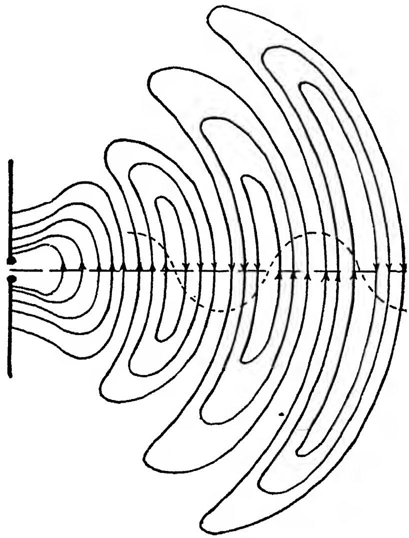


FIG. 208.

corresponding energy reappearing as sound, light, and heat, but if the discharge is sufficiently rapid, that is, if the ends of the tube move quickly enough, the ends of some may cross, forming closed loops as is shown for a single tube in (b), Fig. 207. As the ends continue to move, the condenser being charged in the reverse sense, the loops break off, the original tube giving rise to a condition represented in (c), Fig. 207. Since in reality many tubes will

take part in such a process, at the end of the reversal, the state of affairs may be represented somewhat as in (d), Fig. 207.

After several to-and-fro oscillations we may picture the condition somewhat as in Fig. 208. We now see that, as a result of this rapid oscillatory discharge, in the region surrounding the oscillator, at any instant we have places with the electric intensity in one direction alternating with others where the electric intensity is in the opposite direction. Moreover, it is not difficult to see that, if we fix our attention on any given place, as time goes on the electric intensity at that place will periodically change. *In other words we have the periodicity in time and in space which, as we have seen in Chapter II, characterizes a wave-motion*, and we may represent such waves at any instant by the dotted curve shown in Fig. 208 depicting the changes in electric intensity from point to point in the surrounding medium. To complete the picture, the magnetic lines associated with and running at right angles to the moving Faraday tubes should also be represented. While for simplicity these are omitted, it should not be forgotten that in electromagnetic waves we have to do with periodic changes in both electric and magnetic intensities, that these vector quantities are at right angles to each other, and that each is at right angles to the direction of propagation — in other words, that we are dealing with transverse vibrations.

Now Hertz took such an electric arrangement and by means of a detector proved the actual existence of electromagnetic waves. Moreover, he showed, experimentally, not only that such waves had many of the properties of light waves, but also that they travelled with the velocity predicted by Maxwell. They could be reflected; they could be refracted and the value of the index determined; they were polarized when obtained from such an oscillator, and they gave rise to interference phenomena. The work of Hertz thus completely established the general truth of the electromagnetic theory of light. *Light waves and waves generated electrically are of the same nature.* Many of their properties are, of course, totally different. For example, some electrical waves will readily pass through a brick wall, light waves will not.

Differences in properties arise from the very marked differences in wave-length, but the essential nature of both is the same; for both are electromagnetic periodic disturbances travelling at a fixed speed in ether.

149. Range of Electromagnetic Waves. — In Chapter VII it has been pointed out that, in addition to visible light waves, there is an infra-red region of longer, and an ultra-violet region of shorter, invisible optical wave-lengths. In days when radio is a household word, the reader will scarcely need to be told that since the time of Hertz an enormous amount of work has been done in connection with electrical waves, and that they can be generated with lengths ranging from miles to millimetres. We thus have short electrical and long optical waves. How near do the two regions approach? The answer is provided by a recent investigation of Nichols and Tear (United States) who obtained from an electric oscillator wave-lengths as short as 0.22 mm. while, at the same time, by suitably screening a mercury arc, they found infra-red waves as long as 0.4 mm. There is then no gap between the longest infra-red radiations and the shortest electrical waves. We thus have a continuous range of electromagnetic radiations extending from electric waves almost as long as we wish to have them, through wireless and Hertzian waves until we reach the infra-red; then visible light, followed by the shorter ultra-violet region in which we have seen wave-lengths as short as 40 A have been measured. But the family does not end there, for *X-rays* and *gamma rays* of radium, and possibly the so-called recently discovered *cosmic rays* extend the range much further. While it is not the purpose of this book to give a detailed discussion of all kinds of electromagnetic waves, a brief reference will be made to those we have just enumerated.

150. X-rays. — When electrons moving at high speed strike a metal surface, X-rays originate. Suppose we are provided with a highly exhausted vacuum tube, in which is inserted a filament F, Fig. 209, acting as one electrode, and a target T of solid tungsten

metal, acting as a second electrode. If the filament is heated to incandescence by means of an electric current, while at the same time a big potential difference exists between *F* and *T*, *T* being positive, the electrons emitted thermionically from the filament are driven across the tube at high speed and strike the target. An invisible radiation, called sometimes Roentgen rays, after their discoverer, sometimes X-rays, because of their original unknown nature, then fills the whole region in front of the target and to some extent behind it. These rays, or some of them, pass

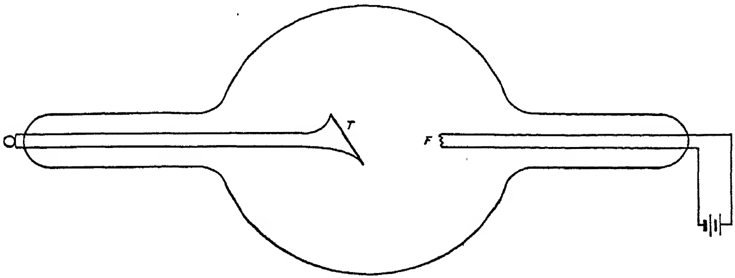


FIG. 209.

through wood as readily as sunlight passes through window glass; they excite fluorescence in certain substances; they ionize air; they affect a photographic plate; they discolour certain salts; and they have profound physiological effects. For many years after their discovery in 1896 their nature was uncertain, but when in 1912 it was shown by Laue, Friedrich, and Knipping (Germany), that they give rise to interference and diffraction phenomena, it was conclusively shown that they too are electromagnetic waves, but with wave-lengths much shorter than visible light waves. Indeed, it was the shortness of the wave-lengths which was the cause of the negative results obtained by the early investigators who sought to show the existence of a wave phenomenon like refraction.

Laue, Friedrich, and Knipping, Bragg and Bragg in England, and their many successors showed that the natural arrangement of

atoms in planes in crystals act towards X-rays much as a diffraction grating does towards light, permitting an accurate measurement of wave-lengths. In the rapid development which followed, it was shown that X-rays is as general a term as light, that we have to do with wave-lengths as short as 0.06 angstrom and as long (in special circumstances) as several hundred angstroms. Different elements, moreover, may emit characteristic wave-lengths just as they do when optically excited, although the beam from an ordinary target consists of a continuous range over a region which depends on the voltage across the tube. Again, just as there are marked differences in the properties of the various groups of electromagnetic waves, so in the single group to which the name X-rays is given, the properties vary markedly with the wave-length. Thus rays as long as 20 angstroms are so feebly penetrating that they do not pass through the walls of a thin glass vessel, while the shortest X-rays will pass through tolerably thick layers of metal. In radiology, the range of wave-lengths utilized ranges only from 0.3 angstrom (very soft rays) to 0.06 angstrom (very hard rays).

From 1912 until possibly 1925, measurements of X-ray wave-lengths were all made either by the use of a crystal as a grating, or, in the case of those of the order of 100 angstroms, by indirect methods. Since that time, however, it has been shown by Compton (United States), Thibaud (France), and others that if the beam of such rays strikes an ordinary ruled diffraction grating at glancing incidence, X-ray diffraction spectra are obtained from which the longer X-ray wave-lengths may be accurately measured. Thus Thibaud has recorded wave-lengths of magnitudes (in angstroms) such as 23.8, 31.8, 44.9, 68.0, 54.9, and as long as 144. Again Osgood, working in Compton's laboratory at the University of Chicago, by the use of a concave glass grating, has measured X-rays of lengths ranging as long as 211A. Now it will be recalled that with an optical source a wave-length as short as 40 A has been measured. It will be seen then that there is now no gap between the shortest extreme ultra-violet wave-lengths and the longest X-rays. Indeed, in many respects, X-rays may be con-

sidered to be a branch of optics. A striking illustration of this is found in the comparison of two grating spectrographs used by Thibaud. The two instruments are essentially similar, but one is equipped with a condensed spark chamber, the other with a metallic X-ray tube. With the first, optical wave-lengths are measured, with the second, very long X-rays. An intelligent study of the special problems arising in the field of X-rays, however, demands a book in itself, and the writer considers it advisable to limit the scope of this book to purely optical problems.

151. Gamma Rays. — The range of electromagnetic waves does not end with the shortest X-rays, for gamma rays of radium bring us into a still shorter region. A few substances, of which radium is a notable example, have the property of *spontaneously* emitting rays which, like X-rays in some respects, ionize a gas, excite fluorescence, and affect a photographic plate. Such substances are said to be radioactive. Analysis of the rays emitted shows that they can be divided into three groups, (1) *alpha* rays, positively charged particles, each of which has a mass equal to that of a helium atom; (2) *beta* rays, negatively charged particles, with mass equal to that of an electron (although depending on the velocity); (3) *gamma* rays, another kind of electromagnetic radiation. The wave-lengths encountered in this region are still shorter than X-rays, the shortest known being about 0.006 angstrom As might be expected, such rays are extremely penetrating.

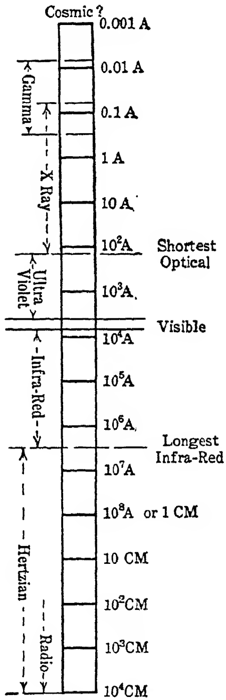


FIG. 210.

152. Cosmic Rays. — In 1903 it was shown by McLennan and Burton, Canadian physicists, that there must exist in the atmosphere a radiation of unknown origin sufficiently penetrating to pass through a layer of water 30 feet thick. This conclusion was in agreement with results obtained by Rutherford and Cooke at Montreal, Canada, and it was confirmed by European physicists, such as Wulf, Gockel, Hess, Kohlhoerster, and others. Observations made on mountain tops, in valleys, in balloons, and below the surfaces of lakes clearly showed that the atmosphere is traversed by a radiation so penetrating that a thickness of several inches of lead will not completely cut it off. As the intensity of this radiation has been found to increase with increasing altitude, its origin cannot be terrestrial. The radiation, whatever it is, comes from outer space — hence the modern name *cosmic rays.*

In recent years the search for the nature as well as for the origin of these rays has been an important field of investigation, and numerous researches, some experimental, some theoretical, have been carried out. In the United States a large number of scientists have co-operated in extensive investigations under the leadership of Millikan and another group under A. Compton, and in Europe the pioneer work has been continued by Regener, Clay, Bothe, Corlin, Piccard, and other observers. In many parts of the world, as well as in stratosphere balloons, observations have been made of changes in the intensity of these rays with altitude and with latitude. Numerous papers have been published and yet, in 1935, questions concerning the nature and the origin of cosmic rays have not yet been fully answered. Just as in radioactive substances we have an emission of both high speed particles and electromagnetic waves, so in cosmic rays we may have to deal with either or both of these kinds of radiation. There is abundant evidence that at least part of the rays is of a particle nature, and it is also certain that, if another part is electromagnetic, the wave-lengths are shorter even than gamma rays. Indeed, wave-lengths as short as 0.00008 angstrom have been estimated.

Although a detailed discussion of the origin of cosmic rays is out of place in a book on optics, it is not amiss to quote Comp-

ton's statement that "nearly all the incoming rays (into the stratosphere) are affected by the earth's magnetic field, and are hence electrically charged." (In the discussion of the Zeeman effect on page 362 reference is made to the magnitude of the force acting on an electron moving in a magnetic field.) Moreover, the combination of intensity observations and calculations based on the earth's magnetic field gives evidence of the existence of three groups of particles. In the order of increasing penetration these include (1) *alpha* particles, (2) electrons (positive and negative), (3) protons.

153. In Table XIII will be found approximate values for the various regions of electromagnetic waves, while the same information is given in a more graphical form in Fig. 210. In this figure a logarithmic scale is used for convenience.

TABLE XIII

Range of Electromagnetic Wave-Lengths

Electric and Hertzian		20 + miles	to	0.22 mm.
Infra-red		0.4 mm.	to	0.0007 mm.
Visible Light		0.0007 mm.	to	0.0004 mm.
	or	7000 A	to	4000 A.
Ultra-violet and				
Extreme ultra-violet		4000 A	to	40 A.
X-rays		500 A	to	0.06 A.
X-rays in radiology		0.3 A	to	0.06 A.
Gamma		0.3 A	to	0.006 A.
Cosmic Rays ?			?	

154. Electric Vector the Light Vector. — In section 148 it was pointed out that in an electromagnetic wave-train we have to do with a periodicity in both the electric and the magnetic intensity, and that the directions of these two vector quantities are at right angles. Again, in the chapter on polarization, we have seen that in plane polarized light, the vibrations are confined

to a single definite plane. Now, on the electromagnetic theory, if the electric vibrations are at all places in the same plane, the magnetic vibrations must be in a plane at right angles to this one. Is it then the periodic changes in the electric intensity we are to think of as light vibrations, or the corresponding changes in the magnetic intensity? The answer to that question is provided when the reflection of light from the surface of a transparent medium like glass is considered on the electromagnetic theory. In the mathematical treatment, which is beyond the scope of this book, it is shown that when reflection takes place at the polarizing angle, the electric vibrations are normal to the plane of incidence. But, in Chapter XIII, we saw that a consideration of the phenomenon of scattering showed us that light vibrations must be perpendicular to the plane of incidence in polarized reflected light. It follows, therefore, that it is the electric vibrations which are the light vibrations.

Experimental confirmation of the truth of this was provided by Wiener in an experiment which showed the existence of stationary light waves in a thin film of silver chloride in collodium, when direct and reflected beams of light meet. Since at a node the sensitive film is not acted on, while at an antinode there is a reaction, the film contains layers alternately black and light (regions where the silver is separated out alternately with those where it is not). Now when electromagnetic waves are reflected and stationary waves result, the magnetic and electric nodes are a quarter of a wave-length apart, and Wiener's experiment showed that the position of the nodal regions corresponded to the predicted position of the electric nodal places, not the magnetic.

155. Theory of Dispersion. — Hitherto, although much has been said about the phenomenon of dispersion, no attempt has been made to explain why the velocity of light in transparent bodies varies with the frequency. The general reason lies in the fact that in such cases we are dealing with the effect of particles of matter on the propagation of the wave disturbance through the ether in which the matter may be considered to be imbedded.

Just what effect this is, we shall briefly consider from the point of view of the electromagnetic theory.

In the first place, the student should recall the modern view that an atom consists of a positively charged nucleus surrounded by one or more negatively charged electrons. When an atom is a constituent part of a body *which is a conductor of electricity*, and an electric force is applied, one or more of the electrons may readily be passed on to a neighboring atom, thus giving rise to an electric current. On the other hand, *if the body is a dielectric*, and that is the kind in which we are here interested, the electric force causes only a displacement of the electron. If the electric force is constant, the electron comes to rest when the magnitude of the *restoring forces*, called into play as a result of its displacement, is equal and opposite to the electric force. If such a displacement is caused by the application of a force which is at once removed, the restoring forces may be considered to cause the electron to vibrate to and fro much as a released mass at the end of a stretched spring executes a periodic motion. The period of an electron, when so displaced and left free to vibrate, represents a *natural frequency* of vibration. In a dielectric such as glass, therefore, we can think of certain electrons in the constituent atoms as having natural frequencies of oscillation, provided they have suffered, for some reason, an initial displacement. In the present work we shall confine our attention to the case where there is only one such natural frequency for a given dielectric. (The theory can readily be extended to include the more general case of several such frequencies.)

Consider now what happens when a rapidly *alternating* electric force is applied to such an electron, a condition which occurs on the electromagnetic theory, when the dielectric is traversed by a beam of light. The electron is then forced to execute oscillations with the period of the applied force, and consequently the amplitude of these forced oscillations depends on how nearly the frequency of the applied force approaches the natural frequency of the electron. If the two are about equal, we have the phenomenon of *resonance*, with which work in sound should have made the

student familiar. In that case, the electron vibrates with a very large amplitude, almost all of the energy of the incident disturbance being then absorbed. If in the exciting beam there is a large range of frequencies (or wave-lengths), we then have an *absorption band* in the spectrum of the emerging light, — a point to which reference has already been made in section 72, Chapter VII. As such absorption represents a loss of energy from the original beam, it is natural to assume, when dealing with the motion of the vibrating electron, that there is a factor akin to friction whose magnitude is proportional to the velocity of the electron. In the case of resonance, that is, of an absorption band, amplitudes and hence velocities are very large, and the frictional losses very marked. If, however, we deal with frequencies well away from the natural frequency of the electron, the amplitude of the forced vibration is small, and therefore also the effect of friction.

When these ideas are put into mathematical form, and Maxwell's equations are extended to include currents due to the movement of the electrons, then, instead of the simple relation $n^2 = k$ deduced above, one obtains the expression

$$n^2 = 1 + \frac{A_1\lambda^2}{\lambda^2 - \lambda_1^2}, \tag{16.02}$$

where n is the index of refraction of the dielectric for wave-length λ, where A_1 is a constant whose value depends on the charge on the electron, on its mass, on its natural frequency, and on the number of electrons in unit volume; and where λ_1 is the wave-length in free ether corresponding to a frequency equal to the natural frequency of the electron.

(If there is more than one natural frequency, additional similar terms must be added to this expression.)

Now let us examine relation (16.02) a little more carefully with reference to one or two cases.

(1) **For very long waves,** that is, when λ is many times greater than λ_1.

In this case, the expression reduces to

$$n^2 = 1 + A_1, \tag{16.03}$$

which corresponds to the $n^2 = k$ discussed above (16.01). Dispersion does not now enter into the question.

(2) When λ is sufficiently long, that the ratio $\frac{\lambda_1}{\lambda}$ is small. For example, if, in a transparent substance like glass, the absorption band (λ_1) is well in the ultra-violet, and we confine ourselves to visible wave-lengths, we approximate to this condition.

Re-writing equation (16.02), we have

$$n^2 = 1 + \frac{A_1\lambda^2}{\lambda^2 - \lambda_1^2}$$

$$= 1 + \frac{A_1}{\left(1 - \frac{\lambda_1^2}{\lambda^2}\right)}$$

$$= 1 + A_1\left(1 - \frac{\lambda_1^2}{\lambda^2}\right)^{-1}$$

$$= 1 + A_1\left(1 + \frac{\lambda_1^2}{\lambda^2}\right),$$

if we neglect higher powers than the second of $\frac{\lambda_1}{\lambda}$, as we may do with little error when its value is small. We have, then

$$n^2 = 1 + A_1 + \frac{A_1\lambda_1^2}{\lambda^2}$$

which we may write

$$n^2 = A + \frac{B}{\lambda^2}, \text{ by putting} \tag{16.04}$$

$$A = 1 + A_1, \quad \text{and} \quad B = A_1\lambda_1^2.$$

In other words, equation (16.02) reduces to the simple Cauchy formula which, as was pointed out in section 59, Chapter VI, applies with a fair degree of accuracy to ordinary cases of dispersion.

(3) Anomalous Dispersion. — When the range of λ's includes λ_1, that is, when we have to do with an absorption band in the region considered.

In this case it will be evident that, as we pass from large values of λ in the direction towards λ_1, the value of the index increases more and more rapidly until, for wave-lengths such that $\lambda^2 - \lambda_1^2$ is a small quantity, the value becomes extremely large. For wave-lengths, therefore, which are just *greater* than that corresponding

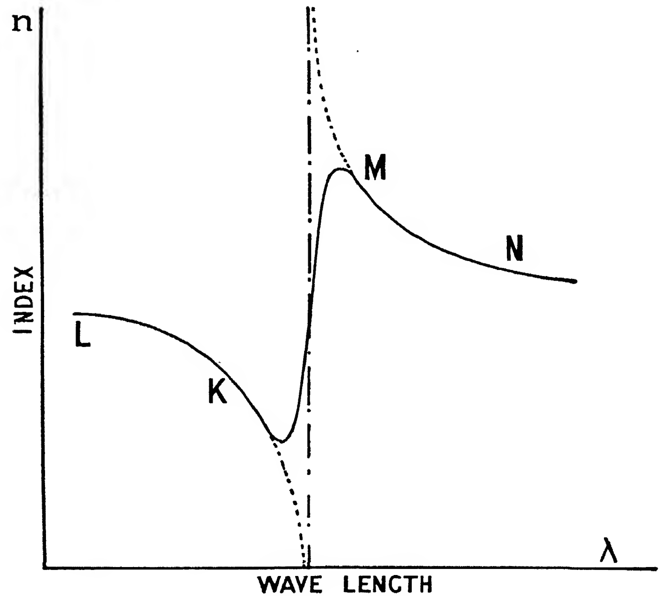


FIG. 211.

to the absorption band, the index is abnormally high, as is roughly indicated by the portion *MN* of the curve in Fig. 211. On the other side of the band, that is, when λ is a little *less* than λ_1, the term $\dfrac{A_1\lambda^2}{\lambda - \lambda_1^2}$ is negative, and we obtain abnormally low values of n which, however, gradually increase as λ becomes still smaller — somewhat as shown by the portion *LK* of Fig. 211. For wave-lengths on one side of an absorption band, therefore, the index should be abnormally high, on the other side, abnormally low.

This means that, if one used a prism of a material which had an absorption band in the green, for example, that wave-lengths beyond the band on the violet side should be refracted *less* than those near the band on the longer wave-length side. As a matter of fact, such cases of so-called *anomalous dispersion* are well-known. Thus a prism made of an alcoholic solution of fuchsine, an aniline dye with an absorption band in the green, refracts violet light less than red, although red, orange and yellow appear in the normal order.

In Chapter VII attention was directed to the selective absorption of sodium vapour, which gives a dark band (strictly speaking, two close together) in the yellow part of the visible spectrum. That sodium vapour exhibits anomalous dispersion should be clear from an inspection of the photograph reproduced in Fig. 5, Plate III. In taking this photograph a glass prism, which refracted a beam of white light into a narrow *horizontal* spectrum, was used in conjunction with an exhausted cylindrical horizontal tube containing pieces of metallic sodium. On strongly heating this tube at its *lower* surface the sodium was vapourized, evidently in such a way that the vapour density was much greater at the bottom of the tube than at the cooler top. The column of vapour then acted as a prism with base at the bottom and refracting edge at the top. In this way what was equivalent to a sodium vapour prism deviated the light vertically, thus making evident, in the manner shown in the photograph, the phenomenon of anomalous dispersion. In spite of the name "anomalous," the student should realize from the above explanation that ordinary dispersion and anomalous are just two particular features of the same general phenomenon.

CHAPTER XVII

THE ORIGIN OF SPECTRA

Preliminary Discussion

156. Throughout this book thus far, we have dealt almost entirely with questions relating primarily to the propagation of light, and but little has been said about processes of emission and absorption. We wish now to see, in the light of the facts and theories so far presented, what can be said concerning the ultimate source or origin from which light waves emanate.

(1) **The ultimate source, whatever it is, must in some way be related to the atom.** Consider the evidence.

(a) The mere fact that the frequency of any kind of visible light is so high *suggests* that the dimensions of the "vibrator" are extremely small in comparison with those of ordinary pieces of matter. Thus, for $\lambda = 5 \times 10^{-5}$ cm., since the velocity of light $=$ 3×10^{10} cm. per second, the frequency $= \dfrac{3 \cdot 10^{10} \cdot 10^{5}}{5}$ or about 6×10^{14} per second. If, therefore, anything material is oscillating, mechanical considerations suggest that it is extremely small, and it is natural to think of the atom as a possible source.

If the electromagnetic theory is accepted, the shortness of the waves suggests the same thing, for, in generating the very short electric waves which closed the gap between the infra-red and the electric region, oscillators of minute dimensions are necessary. (Spark lengths as short as a few thousandths of a millimetre were used.) If, then, electric waves as short as those encountered in light are generated, it is not improbable that the oscillator is of atomic dimensions.

(b) This view is supported by the fact of spectrum analysis. When just a trace of a compound such as sodium chloride is placed

in a bunsen burner, lines characteristic of the metallic element are emitted — a fact which is to be expected if the resulting radiation is from atoms. (It may not be amiss to state here that later we shall present evidence that we also obtain light as a result of radiation from molecules. This, however, in no way contradicts the general proposition we are discussing.)

(c) The hypothesis that the ultimate source has to do with the atom provides a ready explanation of many observed facts, when it is considered with reference to the translatory motion of atoms and molecules postulated by the kinetic theory of gases. To understand the nature of this evidence the student should recall the work of Chapter VI, where it was shown that in an ordinary prism spectrograph, each spectral line is an image of the illuminated slit of the instrument. It follows that the narrower the slit, the narrower the observed spectral line, a conclusion which is in general accord with experiment. If, however, an extremely narrow slit is used, the observed width of the line is found to vary with certain factors. For example, in the case of a luminous gas or vapour, the width of the spectral line varies with the pressure, being narrower at low pressures. On the other hand, no matter how low the pressure, a certain limiting width is ultimately reached. Again, if a vacuum tube containing a luminous gas such as helium is plunged into liquid air, it has been shown by the use of high-power spectroscopes that the spectral line is narrower than at room temperatures, or, more accurately, *that the light corresponding to a single spectral line is more nearly absolutely monochromatic at the low temperature than at the high.*

Finally, if two vacuum tubes are taken, one with luminous helium of low atomic weight, the other with luminous mercury vapour, of high atomic weight, it can be shown that a single spectral line is narrower, that is, more nearly monochromatic, in the case of mercury than of helium.

All these facts find a ready explanation in terms of the kinetic theory of gases. According to it, in a gas or vapour, the atoms (or molecules) are moving about in an irregular " zig-zag " fashion with velocities grouped about a mean square velocity whose

magnitude is greater, the higher the temperature, and the smaller the mass of the atom. Suppose now that a single atom, or a part of one, is radiating light of a single definite frequency, as well as a few million other atoms of the same kind. Because of the motion of translation postulated by the kinetic theory, some of these atoms will be moving towards the slit of a spectrograph receiving the light, others away from it, others along a line at right angles to the axis of the collimator of the spectrograph, and still others in other directions. If the Doppler effect is evident, as it should be under such conditions, it follows that the instrument will record a small range of wave-lengths grouped around that particular wave-length corresponding to the frequency in the source. In other words, no matter how narrow the slit, some finite width is to be expected — as is actually the case.

Again, if the temperature of the luminous gas is suddenly lowered, the velocities of the atoms become less, hence the changes in wave-lengths arising from the Doppler effect become less, and the spectral line becomes more nearly monochromatic. The helium tube in liquid air should radiate light more monochromatic than at room temperature.

If two groups of atoms of different elements, such as helium and mercury, are radiating light while at the same temperature, the heavier mercury atoms, moving more slowly, emit light more nearly monochromatic than the lighter helium atoms.

Finally, if a luminous gas is at higher and higher pressures, it can readily be seen that collisions between atoms become more and more frequent until ultimately an individual atom, because of frequent "shocks" and the resulting short mean free path is no longer left isolated sufficiently long to emit monochromatic radiation. In this connection we may recall that an incandescent solid emits a continuous spectrum.

We conclude, then, that many facts find a ready interpretation, if the source of light is closely related to the atom.

(2) **The ultimate source must be electrical in nature.**

(a) If the electromagnetic theory is accepted, this conclusion is inevitable. Electric waves demand an electric source.

(b) This view is supported by the Stark effect, a luminous source being influenced in a marked degree by a strong electric field, and

(c) by the Zeeman effect, where the field is magnetic.

157. If, then, the source is electric, and if it is intimately connected with the atom, what is it? To that question, a very natural answer suggests itself. Why not the electron, a constituent part of every atom, as well as a negatively charged particle? Certainly it is easy to form a *qualitative* picture in which an electron, displaced from its normal position by some external agency, is acted on by a restoring force as a result of which it vibrates to and fro, radiating light; or, a somewhat different picture, in which the electron rapidly revolving around the positive nucleus, much as the planets encircle the sun, radiates waves whose length is determined by the period of revolution. According to the electromagnetic theory, whenever charges are accelerated, as in either of these two pictures, radiation results. "Acceleration of electric charges is the only known mode of originating ether disturbances" — so wrote Sir Oliver Lodge at the beginning of this century. It was natural, then, that such pictures should be examined quantitatively to see if observed facts of light could be deduced from such hypothetical sources.

Before pointing out the objections to accepting either of these plausible pictures, let us consider the explanation of the Zeeman effect from the above point of view, an explanation which, although presented in an elementary fashion, embodies ideas somewhat similar to those originally given by Lorentz shortly after the discovery of this phenomenon in 1896. To understand what follows, the student needs to recall the "motor principle," the law that a portion of a circuit (carrying a current), which is free to move and which lies in a magnetic field, is acted on by a force, at right angles both to the direction of the current and to that of the field, of a magnitude directly proportional both to the strength of the magnetic field and to that of the current. If we have not an ordinary *wire* circuit but a current due to a flight of charged

particles, each of mass m and each bearing a charge e, moving with velocity v, then the force acting on the particle is equal to Hev, where H is the intensity of the magnetic field. An electron, therefore, which rotates in an orbit of radius r about the nucleus of an atom with positive charge E, in a plane perpendicular to the lines of a magnetic field in which the atom lies, is acted on by the force Hev in addition to the electrostatic force of attraction Ee/r^2. Moreover, the direction of this force is either towards or away from the centre, depending on the sense of rotation. If the force Hev is towards the centre, the resultant force is greater than when no magnetic field is present, and consequently the rotating electron moves at a higher speed. When the force is away from the centre, the opposite is the case and the speed is slower.

Now, without going into dynamical details, it can be shown that the most general motion of a particle like an electron about a centre of attraction is dynamically equivalent to the superposition of a linear vibration along any specified direction, and two equal and opposite circular motions in a plane at right angles to this direction. If, then, we have a few million million atoms in a magnetic field, we can think of one group with electrons oscillating along the lines of force (the specified direction), another group with electrons executing circular oscillations in one sense, in planes perpendicular to the lines, and a third group with circular motions in the opposite sense. The first motion is unaffected by the magnetic field (being along the lines of force), but the circular motions, originally of equal periods, have their periods altered, one slightly decreased, the other slightly increased. If, then, an observer or an instrument is receiving light in a direction perpendicular to the lines of force, instead of one beam of approximately monochromatic light (a single spectral line), three should be received, one with wave-length the same as that transmitted in the absence of the magnetic field, a second with wave-length slightly greater, and a third with wave-length slightly less than the first. In the so-called normal Zeeman effect this is exactly what is observed, the original spectral line being replaced by three others as is nicely shown in

Fig. 3, Plate III, a reproduction of a photograph given the writer (along with those shown in Fig. 4, Plate III) by Dr. J. C. McLennan of the University of Toronto. It should be understood that the change in wave-length is very slight, actually being equal to 0.041 A in the case illustrated. The effect would not be visible, therefore, in an ordinary prism spectograph, an instrument of very high dispersion and resolving power being necessary.

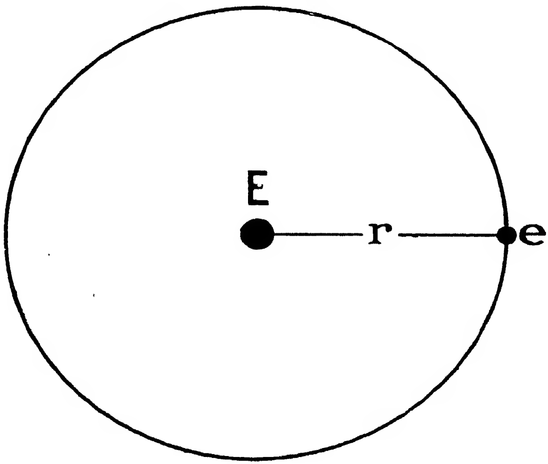


FIG. 212.

A little reflection should make it clear that all three wave-lengths are linearly polarized, the vibration plane of the central line being along the lines of force, that of the other two at right angles to this direction. This can be readily shown by observation through a Nicol prism or double-image prism.

If the instrument is directed along the lines of force, there is no light received corresponding to the linear motion, for the vibrations are longitudinal with respect to this direction. Corresponding to the two circular motions, however, two disturbances are received, one with wave-length slightly greater, the other slightly less than the normal, the two beams being circularly polarized in opposite senses.

The Zeeman effect for longitudinal observation is nicely shown in Fig. 4, Plate III, where the upper photograph has been taken with field on, the lower with field off. Note that the original single line is now replaced by a pair.

When the above ideas are incorporated into equations of motion, an important result follows. Suppose, as depicted in Fig. 212, an electron of mass m and charge e rotates in a circular orbit of radius r about a nucleus of charge E, no magnetic field being present. If v is the linear velocity of the electron and ω its angular velocity, then by ordinary mechanics

$$\frac{mv^2}{r} = mr\omega^2 = \frac{Ee}{r^2}.$$

When a magnetic field of intensity H gauss is applied we have the additional force Hev, directed, for motion in one sense, towards the centre; for motion in the opposite sense, away from the centre. If ω_1 is the angular velocity for the one case, ω_2 for the other, we have now the equations

$$\begin{aligned} mr\omega_1^2 &= mr\omega^2 + Hev \\ &= mr\omega^2 + Her\omega_1 \end{aligned}$$

and

$$mr\omega_2^2 = mr\omega^2 - Her\omega_2.$$

These two equations reduce to

$$\omega_1^2 = \omega^2 + H\frac{e}{m}\omega_1,$$

$$\omega_2^2 = \omega^2 - H\frac{e}{m}\omega_2$$

Subtracting we have

$$\omega_1^2 - \omega_2^2 = H\frac{e}{m}(\omega_1 + \omega_2)$$

or

$$\omega_1 - \omega_2 = H\frac{e}{m}.$$

Now *if* the frequencies of the light radiated are equal to the orbital frequencies, and λ_1, λ, λ_2 are the corresponding wave-lengths, we have

$$c = \frac{\omega_1\lambda_1}{2\pi}, \quad c = \frac{\omega\lambda}{2\pi}, \quad \text{and } c = \frac{\omega_2\lambda_2}{2\pi},$$

where c = velocity of light.

Hence we can write

$$\frac{2\pi c}{\lambda_1} - \frac{2\pi c}{\lambda_2} = H\frac{e}{m}$$

or

$$\frac{\lambda_2 - \lambda_1}{\lambda_1\lambda_2} = \frac{H}{2\pi c}\cdot\frac{e}{m}.$$

Since $\Delta\lambda$, the difference in wave-length between λ_1 and λ (or λ_2 and λ) is small, we may write

$$\frac{2\Delta\lambda}{\lambda^2} = \frac{1}{2\pi}\cdot\frac{H}{c}\cdot\frac{e}{m}$$

or

$$\Delta\lambda = \frac{H}{4\pi c}\cdot\frac{e}{m}\cdot\lambda^2 \tag{17.02}$$

If, then, the value of $\Delta\lambda$ is observed for a measured magnetic field, this expression provides us with a means of evaluating e/m. *It is most significant that values obtained in this way agree with those determined by purely electrical methods.* When, therefore, Lorentz * derived (in a more general and somewhat different way) this relation shortly after the discovery of the Zeeman effect, and when this striking agreement for the values of e/m was obtained, it seemed *as if* strong confirmation had been obtained for the truth of the picture we have been presenting concerning the origin of spectra.

* Lorentz made use of a type of atom in which planetary motion is not necessary, for the electrons occupy positions of equilibrium within a sphere of positive electricity. Radiation was considered to take place when an electron, displaced from its equilibrium position by some external agency, vibrated to and fro because of quasi-elastic forces supposed to be proportional to the amount of displacement. The work of Rutherford on scattering of alpha particles showed that this view of the structure of an atom could not be maintained.

158. Let us, however, look at some objections to this view.

(1) Suppose we consider, as we have done, that in the normal (unexcited) state of an atom, an electron is revolving in a fixed orbit around the nucleus, as a planet revolves about the sun. At once we encounter an objection to the electromagnetic theory, for, since in the motion of the electron, we have a case of accelerated motion of an electrical charge, radiation should result *without excitation by any external agency*. But substances in their normal unexcited state do not radiate.

(2) If, disregarding this difficulty, we assume that such a radiation does take place (as we have done in the discussion of the Zeeman effect), we at once encounter another difficulty, for the resulting radiation should give rise to a *continuous* spectrum. As energy is radiated, the total energy of the atomic system (the positive nucleus plus the rotating negative electron), must become less. This means that the electron must be continuously attracted nearer and nearer to the nucleus, hence that the radius of the orbit and the frequency of rotation continuously change. If, therefore, the orbital frequency is the radiation frequency, as is the case on the view we are considering, the spectrum observed should be continuous. But, if there is one feature which characterizes the ordinary spectrum of a luminous gas or vapour more than any other, it is the appearance of well-defined isolated spectral lines.

(3) Suppose, however, that we adopt the picture of the atom used by Lorentz, considering that an electron is not in rotation but is in some static position of equilibrium as a result of all the forces acting on it. We may then assume that radiation results only when, because of shocks or collisions provided by an external agency, the electron is displaced from its position of equilibrium, radiation taking place as it executes vibrations about this position. When mathematical analysis of such a condition is made, it can be shown, as was done in 1906 by the late Lord Rayleigh, that relations *not* in agreement with the observed laws of spectral series are obtained. In all such calculated relations the square of the frequencies appears, not the simple first power as in series relations.

There are serious difficulties, therefore, when it comes to dealing with the mode of production of light waves, whether it is assumed that radiation results from the orbital motion of an electron, or from the vibrations of an electron displaced from an equilibrium position. In spite of the wonderful success of the electromagnetic theory in so far as questions relating to the propagation of light are concerned, the ablest mathematical physicists at the close of the nineteenth century and the beginning of the twentieth were unable, *as long as they adhered rigidly to the dynamical principles involved*, to find a satisfactory picture of the origin of spectra.

But, it may be asked, what about the striking agreement in the value of e/m obtained from the Zeeman effect compared with those by other methods? A two-fold answer may be given to that question. In the first place, the Zeeman effect is by no means as simple a phenomenon as has thus far been represented. The majority of spectral lines do not give rise to the characteristic group of three explained, apparently so simply, on the view discussed above. The pattern for most lines is more complex, sometimes much more so, and finds no such ready explanation.

In the second place, the quantum theory, a discussion of which will be given in the next chapter, not only is able to provide the same expression for the normal group of three but also to account for the more complex so-called anomalous Zeeman effect observed with other lines.

We conclude, then, that, while there must be much truth in many of the ideas put forward in this chapter, in particular, that the source of radiation is intimately related to the electron, a satisfactory explanation of the origin of spectra cannot be found, — at any rate, has never been found — in any combination of the modern atom with the electromagnetic theory of light.

CHAPTER XVIII

THE QUANTUM THEORY AND ORIGIN OF SPECTRA

159. Birth of the Quantum Theory. — To understand something of the way in which this very important theory originated, the reader is asked first to recall the nature of the spectrum of an incandescent solid. If the light radiated from the filament of a lamp is analyzed by a prism, a continuous spectrum is obtained. By means of an instrument such as a bolometer or a thermopile, a quantitative examination may be made of the way in which the energy is distributed throughout the whole range of wave-lengths. When this is done, curves somewhat similar to those given in Fig. 213 are obtained. Thus, for a particular temperature of the source, there is always a certain wave-length, λ_m we shall designate it, for which the intensity of radiation is a maximum. Moreover, as the graphs of Fig. 213 show, the higher the temperature, the smaller the value of λ_m.

Now no scientist is content simply with the obtaining of an experimental curve. His search is always for causes and for origins, and when curves such as those of Fig. 213 are obtained, the next step is to deduce *theoretically* from fundamental considerations a law which is in agreement with the observed facts. In the case under discussion, it will not do to obtain the spectral distribution energy curve for any incandescent solid, for, even at the same temperature, in general, incandescent solids do not give identical curves. The radiation depends on surface conditions, as well as on temperature. There is, however, one kind of luminous solid which emits radiation whose spectral distribution is a function only of the temperature. Such a substance, called a *black body*, is one which completely absorbs all electromagnetic waves incident on its surface, and theoretical considerations are always made with reference to it. There is, then, in this field of work, a two-fold problem: (1) to examine experimentally with the utmost care the

radiation from a black body; and (2) to see, if, by a consideration of radiation processes, and of the equilibrium state obtaining among various wave-lengths, a law can be deduced in agreement with the experimental results.

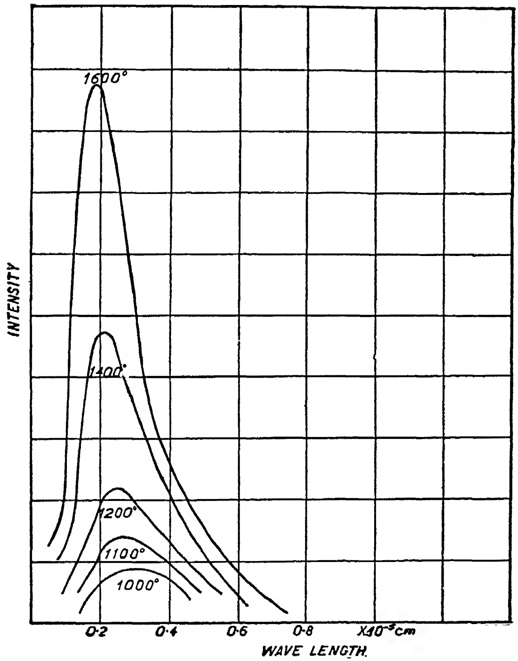


FIG. 213.

To obtain in actual practice a black body, use is made of the fact that in an enclosed cavity kept at constant temperature, the energy is distributed among the various wave-lengths in the same way as in the radiation from an ideal black body. The radiation from such an enclosure is independent of the material of which its walls are made. One such cavity, by means of which accurate measurements have been made, consists of three concentric porcelain cylindrical tubes, the region within the inner of which is

kept at constant temperature by electric heating and the most rigid precautions. The radiation emerging from a small opening from the interior of this inner tube is as close an approximation to black body radiation as human ingenuity has been able to devise. Analysis of the radiation at different temperatures gives curves of the type shown in Fig. 213.

On the theoretical side, the problem has to do with the way in which radiant energy is distributed among the various wave-lengths in an enclosure in which electromagnetic waves are constantly passing and re-passing, while at the walls emission and absorption are constantly taking place. When attempts were made to deduce a law in agreement with the experimental curve grave difficulties were encountered. Indeed, at the outset, it may be stated that so long as no departure was made from the laws of Newtonian mechanics and classical electrodynamics, it proved impossible to deduce a law representing all the facts. Radiation laws were, of course, derived, and two of these it is well to note.

160. Wien's Radiation Law. — Making certain assumptions, Wien proved that

$$E_\lambda = \frac{c_1}{\lambda^5} \cdot e^{-\frac{c_2}{\lambda T}} \tag{18.01}$$

where E_λ is the intensity of radiation for wave-length λ, c_1, c_2 are constants, and T is the absolute temperature. Experiment shows, however, that this law holds only for the region of short wave-lengths.

161. Rayleigh's Radiation Law. — According to this

$$E_\lambda = \frac{c_1 T}{c_2 \lambda^4} \tag{18.02}$$

where c_1, c_2 are constants, T is the absolute temperature. This law holds approximately for long wave-lengths, but not at all for short, as can be seen at once, for according to this relation, the smaller λ, the greater the intensity of radiation, a deduction most certainly not in accordance with the curves of Fig. 213.

162. Planck's Radiation Formula. — In 1900 Max Planck (Germany), a man who had devoted years of study to problems of radiation, showed that a radiation formula in agreement with the facts could be derived if a certain startling assumption were made. According to this *interchanges of energy in processes of emission and absorption take place only in multiples of discrete energy units.* As in his theoretical investigations he made use of electronic vibrators or oscillators, this hypothesis not only introduced the new idea that energy is atomistic but also assumed that accelerated motion of charges does not always give rise to radiation. It may, therefore, properly be described as revolutionary. The unit of energy is called a *quantum*, the magnitude of which, for a frequency ν is equal to $h\nu$, where h is a universal constant, now called Planck's constant. The value of this constant Planck showed to be

$$h = 6.55 \times 10^{-27} \text{ (erg.} \times \text{sec.)}$$

The radiation formula which he derived is given by

$$E_\lambda = \frac{hc^3}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{k\lambda T}} - 1}, \qquad (18.03)$$

where c = velocity of light,
T = absolute temperature.
k = a constant, whose value $= \dfrac{R}{N}$,

where R is the absolute gas constant, and N is Avogadro's number. In passing it may be noted that, taking the value of R from other work, and the value of k from radiation measurements, Planck obtained a value for N in excellent agreement with values obtained by very different methods.

Planck's radiation formula has been proved by the most careful observations to represent the facts accurately. Moreover, it can readily be shown that, for small values of λ, it reduces to Wien's formula, for large values, to Rayleigh's. It is the one and only relation, therefore, which fits the facts, and it is derived by assuming quanta of energy, and that an electromagnetic oscillator may "oscillate" without radiating.

163. Photo-electric Effect and Quantum Theory. — Although the birth of the quantum theory coincided with the opening of the twentieth century, other evidence in support of it had to be established before the theory became of outstanding importance. In 1905, Einstein (Germany) suggested that the idea of quanta need not be restricted to discontinuity in emission and absorption processes, but that it was quite possible that light might be propagated in little bundles of energy (light quanta), each of which maintained its identity throughout its path. It was not, however, until the photo-electric phenomenon discovered by Hertz in 1889 was subjected to a careful quantitative examination by Lenard (Germany), Hughes (Great Britain), Millikan (United States), and others that evidence became overwhelming that *certain* facts could only be explained in terms of Planck's unit of energy $h\nu$. Let us look at some of these facts.

When an insulated metallic plate, negatively charged, is illuminated by light of suitable wave-lengths, it loses its charge because of a photo-electric emission of electrons. Careful investigations concerning the relations between the kind of light the velocity, and the number of electrons emitted, have established the following laws:

(1) A certain critical wave-length is necessary before the photo-electric emission takes place at all. To illustrate, for some metals, red light is ineffective, while with ultra-violet light, the effect is marked. The incident light, therefore, must contain wave-lengths below a critical value.

(2) While the total number of electrons emitted per second varies with the intensity of the incident light, *the maximum velocity with which an electron emerges from the surface of the metal, for a given kind of light, is independent of the intensity.* This very important fact means that when a source of light is moved from a distance of 1 ft. to 10 ft. from the metal surface, there is no difference in the maximum velocity with which the electrons emerge.

(3) In the region of effective wave-lengths, this maximum velocity depends directly on the frequency of the incident light. In other words, the shorter λ the wave-length (or the higher the

frequency), the greater the velocity. In symbols this law, sometimes known as Einstein's Law, is then written.

$$E_{\text{max.}} = a\nu - P, \tag{18.04}$$

where $E_{\text{max.}}$ = maximum kinetic energy of an emerging electron, where a is a constant independent of the metal, and where P is another constant whose value depends on the nature of the metal.

Now look at the significance of these laws. On the wave theory, energy is spread uniformly and continuously over a wave-front. Therefore, the amount of energy incident on a given surface area, 1 sq. mm. for example, must become less and less the farther the source of light is removed. Now it is obvious that an electron comes out of an atom because of the energy supplied by the incident light — and yet it emerges with the same energy whether the source is 1 ft. or 100 ft. away. On a continuous wave-front theory it is difficult to explain this fact, but the explanation is at once forthcoming, if light energy travels in discontinuous bundles, that is, in light quanta, which may be assumed to travel not unlike the corpuscules of the old corpuscular theory. A quantum theory, according to which energy is distributed over a wave-front anything but continuously, then provides a ready explanation.

What about the magnitude of the energy in a quantum in this case? Experiment and relation (18.04) provide the answer to that question. To understand how this is the case, it is necessary to know that in order to measure the maximum velocity of a photo-electron, an opposing electric field is applied with a potential difference of just such a value as will prevent any electrons leaving the metal. Suppose, then, that the application of a potential difference of V_1 volts keeps electrons from escaping when the incident light has a frequency ν_1, and that V_2 volts are required for a frequency ν_2. Then, remembering that the work done when a particle with electronic charge e falls through a potential difference of V is Ve, we have

$$V_1e = a\nu_1 - P$$

$$V_2e = a\nu_2 - P,$$

from which

$$(V_1 - V_2)e = a(\nu_1 - \nu_2)$$

or

$$a = \frac{(V_1 - V_2)e}{\nu_1 - \nu_2}.$$

The constant a can be evaluated, therefore, from the observed potential values V_1 and V_2, from the frequencies ν_1 and ν_2 of the incident light, and the charge e on an electron. The result of careful measurements made in this way by Millikan and others, gives the value of $a = 6.56 \times 10^{-27}$ (erg. $\times$ sec.), and it is the same for all metals. It is seen, then, that a is in reality Planck's constant h, and that relation (18.04) should be written

$$E_{\text{max.}} = h\nu - P \tag{18.05}$$

The conclusion is obvious. The photo-electric effect receives a ready explanation in terms of a quantum theory, according to which light energy is restricted to multiples of a quantum of magnitude $h\nu$.

164. Quantum Theory and Continuous Spectrum of X-rays. — In the brief reference which has been made to X-rays, it was stated that this type of electromagnetic radiation originates when rapidly moving electrons are suddenly stopped. We may consider then that a portion of the kinetic energy of the electrons is transformed into Roentgen rays, and it is very significant that a quantum relation holds in this interchange of energy. To understand the relation it is necessary to know that, when a beam of X-rays from the target of an X-ray tube is analyzed into its constitutent wave-lengths, it is always found that there is a continuous range with a sharp limit on the short wave-length side. Now, if ν is the frequency corresponding to this shortest wave-length, and V the potential difference through which the beam of electrons incident on the target has fallen, experiment shows that the following relation holds exactly

$$Ve = h\nu \tag{18.06}$$

Since Ve measures the kinetic energy of an electron just before im-

pact, this relation tells us that in the energy exchange, a quantum of radiant energy is of fundamental importance.

165. Quantum Theory and Emission of a Single Spectral Line. — Evidence of a character somewhat related to that pertaining to X-rays is provided by certain experiments in which light emission results from the bombardment of a gas or a vapour by much slower moving electrons. In Chapter VII it was pointed out that, as the speed of the bombarding electrons is gradually increased, more and more spectral lines are emitted, a point illustrated by Fig. 1, Plate II. In *a*, Figure 1, it will again be noted that the spectrum of magnesium consists of a single line, a phenomenon which can be reproduced with other elements. One of the first elements investigated was mercury, in which case it was shown by Franck and Hertz (Germany) that a quantum relation again connects the frequency of the single line emitted with the energy of the colliding electron. (For other elements, the same relation was shown by other observers.)

In mercury, experiment shows that the wave-length 2536 A is alone emitted when the bombarding electron has a velocity acquired by falling through a potential difference of 4.9 volts. Now let us make one or two simple calculations, being given that $e = 4.77 \times 10^{-10}$ absolute electrostatic units of quantity, that 300 volts = 1 absolute electrostatic unit of potential difference.

For wave-length 2536 A, we have

$$\nu = \frac{c}{\lambda} = \frac{3 \times 10^{10}}{2536 \times 10^{-8}},$$

where c is the velocity of light in *vacuo*.

Therefore, the corresponding quantum of energy is

$$h\nu = \frac{6.56 \times 10^{-27} \times 3 \times 10^{10}}{2536 \times 10^{-8}}$$

$$= 7.76 \times 10^{-12} \text{ ergs.}$$

If, now, we assume the relation (18.06), that is, that

$$Ve = h\nu,$$

we have

$$V = \frac{7{\cdot}76 \times 10^{-12}}{4{\cdot}77 \times 10^{-10}} \text{ absolute units}$$

$$\frac{300 \times 7{\cdot}76 \times 10^{-12}}{4{\cdot}77 \times 10^{-10}} = 4{\cdot}[illegible] \text{ volts.}$$

We see, then, that on the assumption that a quantum of energy is radiated when the colliding electron loses an equivalent amount of kinetic energy, there is agreement between observed and calculated values.

166. Bohr's Theory and Origin of Spectra. — The most striking evidence in favor of a quantum theory was given by the Danish physicist Niels Bohr, who in 1913 deduced from theoretical considerations the Balmer series relation for hydrogen spectral lines. As a foundation on which to base his views, Bohr accepted, (1) the planetary view of the atom which Rutherford's work on alpha particles necessitates, and (2) the idea of quanta of radiant energy. Since a planetary electron means an accelerated charge, the difficulties given at the end of the last chapter had to be met. This was done by making the bold assumption that an electron revolving around the nucleus of an atom could occupy *without radiating* certain definite orbital paths, — non-radiating orbits or *stationary states* we shall call them. Now, while this view is nothing but a flat denial of one aspect of the electromagnetic theory (according to which accelerated charges must radiate), it is quite in accordance with the work of Planck on black body radiation.

In any orbit the ordinary laws of mechanics are supposed to hold, that is, the attraction of the nucleus for the electron provides the necessary central force to keep the latter in its path. In the normal state of a non-luminous gas or vapour, the electron occupies an innermost *stable* orbit. When, however, the gas or vapour is "excited" to luminosity by some external agency, the electron is moved farther away from the nucleus, to do which work must be done against the attraction of the nucleus. The total energy of the atom then increases having values E_1, E_2, E_3, . . . E_n, as the

electron occupies the first (or innermost orbit), the second, the third . . . the nth orbit. According to the ideas we are now presenting, radiation results when, after such an increase in energy has been brought about, an electron drops back from an outer orbit to one nearer the nucleus. The frequency of the radiation is given by the condition (known as Bohr's *frequency condition*)

$$h\nu = W_n - W_k, \tag{18.07}$$

where W_n and W_k are the energy values corresponding to the nth and the kth orbit.

Since $\nu = \frac{c}{\lambda}$ we may write equation (18.07)

$$\frac{1}{\lambda} = \frac{1}{hc}(W_n - W_k) \tag{18.08}$$

On this view, then, *each wave-length corresponds to a definite change in energy values, or in what we may call energy levels* and, as we shall see presently, a spectral series consists of a group of wave-lengths, each one of which results from a drop of the electron from an outer to the same inner orbit. Thus we might have a series consisting of wave-lengths $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \ldots$ where

$$\frac{1}{\lambda_1} = \frac{1}{hc}(W_2 - W_1);\ \frac{1}{\lambda_2} = \frac{1}{hc}(W_3 - W_1);\ \frac{1}{\lambda_3} = \frac{1}{hc}(W_4 - W_1);\ \text{etc.}$$

Again, if the external agency exciting the atom to luminosity is such that the electron is displaced from the innermost orbit to the next, obviously only one subsequent change in energy value is possible; in other words, a single spectral line is emitted.

To fix the definite discrete orbits in which an electron may revolve without radiating, that is, to fix the energy levels, another condition is necessary. This is the so-called *quantum condition* according to which the angular momentum of the atom for the nth non-radiating orbit is $n \cdot \frac{h}{2\pi}$, where n can have only integral values, and h is Planck's constant. In thus restricting the angular momentum to multiples of a fundamental unit involving h, this physical quantity is said to be *quantized*.

Now let us apply these ideas to the hydrogen atom, in which case a single electron with negative charge e revolves in an orbit about a nucleus with positive charge $E = e$. In the simplest case the orbits are circular, and we shall so consider them in this elementary treatment. We shall also assume that, because the mass of the nucleus is over 1800 times that of the electron, it is at rest. Our problem is then to see what quantitative relations can be deduced, accepting the above conditions.

Let a = radius of any orbit,
ω = angular velocity of electron.

Since the central attraction of the nucleus for the electron is $\frac{Ee}{a^2}$, we have, applying ordinary mechanics,

$$ma\omega^2 = \frac{Ee}{a^2} \qquad \text{or} \qquad ma^2\omega^2 = \frac{Ee}{a}. \tag{18.09}$$

Before the frequency condition can be applied, an expression giving the energy for any orbit must be found. Since a portion of the energy is potential, its magnitude must be measured with reference to some arbitrary zero, just as elevations are given with respect to sea-level as zero, or just as the potential energy of a raised weight is expressed with reference to its value at the surface of the earth. The arbitrary zero chosen is the state when the electron has been removed completely away from the attraction of the nucleus. *The potential energy for any orbit, therefore, bears a negative sign and is numerically less the farther the orbit is from the nucleus.*

If then $E_{\text{pot.}}$ = potential energy of the atom, we have (since the potential at a distance a from a charge E is $\frac{E}{a}$)

$$E_{\text{pot.}} = -\frac{Ee}{a} \tag{18.10}$$

Again, the kinetic energy $E_{\text{kin.}}$ is given by

$$\begin{aligned} E_{\text{kin.}} &= \tfrac{1}{2}mv^2 \\ &= \tfrac{1}{2}ma^2\omega^2 \end{aligned} \tag{18.11}$$

$\therefore W$, the total energy $= E_{\text{pot.}} + E_{\text{kin.}}$

$$= -\frac{Ee}{a} + \tfrac{1}{2}ma^2\omega^2$$

$$= -\frac{Ee}{2a}, \text{ from relation (18.09)}$$

or

$$= -\tfrac{1}{2}ma^2\omega^2$$

This expression for the total energy then tells us that the *smaller* a, the radius of an orbit, the larger the (negative) number giving the value of the total energy. This is, of course, because of the zero chosen for the measure of the potential energy.

Next let us apply the quantum condition according to which

$$ma^2\omega \text{ (angular momentum)} = n\cdot\frac{h}{2\pi},$$

if a is the radius of the nth orbit.

Since $ma^2\omega^2 = \frac{eE}{a}$, by relation (18.09), we have at once

$$a\omega = \frac{2\pi eE}{nh} \qquad (18.10)$$

or

$$\frac{eE}{a} = \frac{4\pi^2me^2E^2}{n^2h^2}. \qquad (18.11)$$

Therefore, if W_n = energy for the nth orbit,

$$W_n = -\tfrac{1}{2}\frac{4\pi^2me^2E^2}{n^2h^2} = -\frac{2\pi^2me^2E^2}{n^2h^2}. \qquad (18.12)$$

In passing we note that, by relation (18.11)

$$a = \frac{n^2h^2}{4\pi^2meE}. \qquad (18.13)$$

Since $e = 4.77 \times 10^{-10}$ and $E = e$ for the hydrogen atom, and since $h = 6.55 \times 10^{-27}$,

$$\frac{e}{m} = 5.3 \times 10^{17},$$

(all electrical units being in the electrostatic system), relation (18.13) gives for the first or innermost orbit the value

$$a = 0.53 \times 10^{-8} \text{ cm}.$$

This value is in good agreement with results for the radius of a hydrogen atom estimated in other ways.

Relation (18.13) also tells us that the radii of the orbits increase as the square of the natural numbers.

We are now in a position to apply the frequency condition (18.08)

$$\frac{1}{\lambda} = \frac{1}{hc}(W_n - W_k),$$

where λ is the emitted wave-length corresponding to an electron drop from the nth to the kth orbit, or less pictorially, to a change in the atomic energy of amount $W_n - W_k$. Using the value of W_n given by relation (18.12) this gives us

$$\frac{1}{\lambda} = \frac{1}{hc}\left(-\frac{2\pi^2 me^2E^2}{n^2h^2} + \frac{2\pi^2 me^2E^2}{k^2h^2}\right)$$

$$\frac{2\pi^2 me^2E^2}{h^3c}\left(\frac{1}{k^2} - \frac{1}{n^2}\right) \qquad (18.13)$$

or, using the notation of Chapter VII,

$$\bar{\nu} = \frac{2\pi^2 me^2E^2}{h^3c}\left(\frac{1}{k^2} - \frac{1}{n^2}\right)$$

$$= N\left(\frac{1}{k^2} - \frac{1}{n^2}\right), \quad \text{where } N = \frac{2\pi^2 me^2E^2}{h^3c}. \qquad (18.14)$$

It will at once be noticed that this relation is of exactly the same form as that describing any of the hydrogen series given in Chapter VII. Moreover, if k is taken equal to 2, (and if the value of N proves to be the same as the Rydberg constant), the relation is *identical* with the Balmer series formula. Now, since e, e/m, h and c are all fundamental quantities (whose magnitudes have already been given), and since $E = e$ in the hydrogen atom, the value of N can readily be calculated. On substituting the values of these constants, we find

$$N = \frac{2\pi^2 me^4}{h^3c} = 1.09 \times 10^5 \text{ cm.}^{-1}$$

When this result is compared with the empirical value 109678 given in Chapter VII, the striking agreement will be apparent.

One great achievement of Bohr's theory then lies in the theoretical deduction of empirical series relations. Thus lines of the Lyman series arise when an electron drops from outer orbits to the innermost, or normal orbit, for which $k = 1$. For the emission of the Balmer series lines (those shown in Fig. 1, Plate I), as already indicated, the final position of the electron is the second innermost ($k = 2$), while for the Paschen series the orbit corresponding to the limit of the series is the third ($k = 3$). In Fig. 214 these ideas are rep-

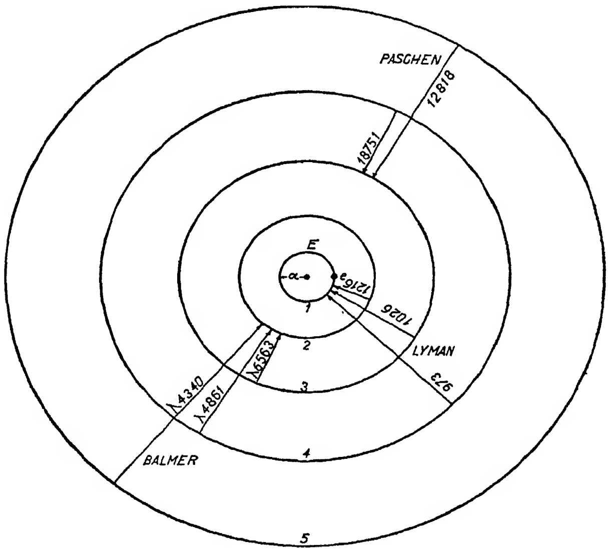


FIG. 214.

resented pictorially. While pictorial representations are often a great aid in grasping certain ideas, it is just as well for the student to be on his guard lest he give to these orbits a reality similar to that of concentric circular grooves in which marbles may roll around a centre. Rather should he think of the orbits as representing definite energy values of the atomic system.

167. Bohr's theory of the origin of spectra had other remark-

able successes to a few of which a brief reference will be made. The first of these has to do with conclusions which follow when the mass of the nucleus is not considered to be infinite. In ordinary hydrogen it is known, from work which has nothing to do with spectroscopy, that M_H the mass of the atom is over eighteen hundred times m that of the electron, or that, to be exact,

$$\frac{M_H}{m} = 1846.$$

Actually, therefore, rotation takes place about the centre of gravity of the nucleus and the electron, a point which of course is very close to the nucleus. When this fact is taken into consideration, and the same principles utilized as in the treatment for a stationary nucleus, it is found that

$$W_n = -\frac{2\pi^2\mu e^2E^2}{n^2h^2}, \text{ where } \mu \quad \frac{m}{1+\dfrac{m}{M_H}} \tag{18.15}$$

It follows that

$$N = \frac{2\pi^2\mu e^2E^2}{h^3c}$$

$$\frac{2\pi^2me^2E^2}{h^3c} \over 1+\dfrac{m}{M_H} \tag{18.16}$$

We see then that the value of N for hydrogen series, N_H we shall describe it, differs *slightly* from $\dfrac{2\pi^2me^4}{h^2c}$, the value previously given, or what we shall now call N_∞, since it was determined on the assumption of infinite mass of the nucleus. Relation (18.16) may now be written

$$N_H = \frac{N_\infty}{1+\dfrac{m}{M_H}} \tag{18.17}$$

Suppose, now, that we have to do with an atom for which a single electron revolves about a nucleus *whose mass is different from that of hydrogen.* Such a system is provided when an atom of

helium, normally consisting of two electrons attracted by a nucleus with a positive charge $E = 2e$, is ionized by the removal of one of the electrons. The resulting atom-ion is, therefore, hydrogen-like, and all the above theoretical considerations may be applied, so that we may write

$$N_{He} = \frac{N\infty}{1 + \dfrac{m}{M_{He}}} \tag{18.18}$$

where N_{He} = value of N for helium,

M_{He} = mass of helium atom.

Combining relation (18.17) with (18.18), we have

$$\frac{N_{He}}{N_H} = \frac{1 + \dfrac{m}{M_H}}{1 + \dfrac{m}{M_{He}}} = \frac{1 + \dfrac{m}{M_H}}{1 + \dfrac{m}{3{\cdot}98 M_H}} \tag{18.19}$$

since $M_{He} = 3{\cdot}98 M_H$.

But from the *empirical* spectral series relations for hydrogen and for helium, we have

$$N_{He} = 109723, \quad \text{and} \quad N_H = 109678.$$

Hence, treating $\frac{M_H}{m}$ as an unknown quantity in relation (18.19), we find

$$\frac{M_H}{m} = 1847,$$

a value almost identical with that given above as obtained by totally different methods.

167A. Heavy Hydrogen and Application of Simple Bohr Theory. — A more recent successful application of Bohr's original elementary theory is connected with the discovery of deuterium or heavy hydrogen. For a number of years it has been known that a large number of elements have two or more different kinds of atoms. The nucleus of each kind has the same positive charge and in each there is the same number of extra-nuclear electrons. But, although the atomic number (see page 153) of each is thus the same, and the chemical properties are often identical, the

masses are different. For example, chlorine, whose atomic weight as determined chemically is 35.46, is really a mixture of two kinds of chlorine, one of atomic weight 35, and the other, 37. Chlorine is said to have two *isotopes*, that is, to be a mixture of two different kinds of atoms which occupy the same place in the periodic table. Mercury, atomic number 80, atomic weight 200.61, has isotopes of atomic weight 196, 198, 199, 200, 201, 202 and 204; lithium, atomic number 3, atomic weight 6.94, has isotopes of atomic weight 6 and 7; and so on for many other elements.

Before 1929 it was considered that hydrogen had only one group of atoms because at that time there was almost perfect agreement between its atomic weight determined by chemical methods (which give only an average value when isotopes are present) and that determined by the mass spectrograph (a purely physical means which permits a measurement of the atomic weight of each isotope which may exist). In that year, however, it was shown (because of the discovery of isotopes in oxygen) that in reality there was a discrepancy of about 2 parts in 10,000 between the chemical and the physical values of the atomic weight of hydrogen. Birge and Menzel (United States) suggested that this discrepancy was due to the presence in ordinary hydrogen of a small amount of an isotope of atomic weight 2. Urey, Brickwedde and Murphy, working at Princeton University, set out to look for such an isotope and in 1932 announced its discovery in a paper published in the *Physical Review* — a paper which already is one of the classics in this era of physics.

From Bohr's elementary theory it is easy to predict the separation of corresponding spectral lines arising from the two kinds of hydrogen. Take, for example, H_α, the first line of the Balmer series. Using relation (7.02), page 165, and (18.17), page 382, we can write

$$\frac{1}{\lambda} \quad \frac{N_\infty}{1 + \dfrac{m}{M_H}} \cdot \left(\frac{1}{2^2} - \frac{1}{3^2}\right).$$

Now if hydrogen has two isotopes with $M_{H^1} = 1$ and $M_{H^2} = 2$, there are two corresponding wave-lengths, H_α^1 and H_α^2, we

PLATE VII

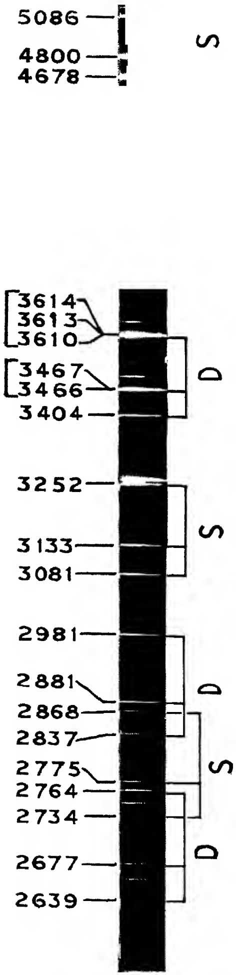


FIG. 1. Cadmium Arc Spectrum. (Quartz Spectrograph)

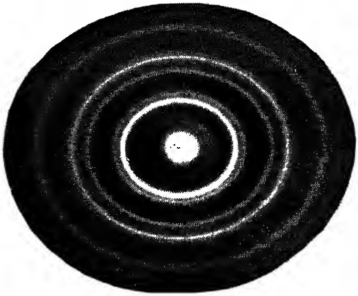


FIG. 2. Electron Diffraction Pattern. (G. P. Thomson)

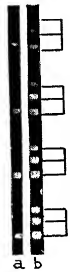


FIG. 3. Fabry-Perot fringes. (a) Line without structure; (b) Line with three components. (E. Gwynne Jones)

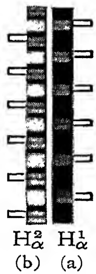


FIG. 4. Fabry-Perot fringes. (a) H^1_α showing doublet structure (b) H^1_α along with H^2_α.

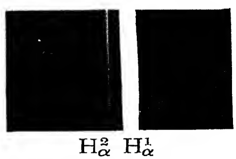


FIG. 5. Grating photograph of H^1_α and H^2_α. (Harris, Jost and Pearse)

shall call them, and it is a matter largely of arithmetic to show that if

$$\lambda \text{ for } H^1_\alpha = 6562.79 \text{ A.}$$

then

$$\lambda \text{ for } H^2_\alpha = 6561.00 \text{ A.}$$

The two wave-lengths, therefore, according to calculation based on Bohr's simple theory, differ by 1.79 angstroms. Using a concave grating, Urey, Brickwedde and Murphy showed that H^1_α was accompanied by the companion line H^2_α separated from it by this amount. Subsequently, means were found, not only of increasing the concentration but also of preparing pure samples of the heavy isotope, to which the name deuterium has been given. (Although deuterium is called an isotope of hydrogen, it is well to note that its properties are by no means identical with those of the lighter isotope.)

Fig. 5, Plate VII, is a reproduction of a grating photograph, reproduced through the kindness of Dr. R. W. B. Pearse, of the Imperial College, London, which shows clearly H^1_α and H^2_α. Reference has already been made to the Fabry-Perot photographs shown in Fig. 4, Plate VII. We may note again, however, that in Fig. 4 (*a*), the set of fringes is due to H^1_α, whereas in Fig. 4 (*b*), the intermediate set arises from H^2_α.

168. Another success of Bohr's theory to which the student's attention is called at this stage has to do with the spectrum of ionized helium. Consider again an atom of this helium, for which, as already pointed out, $E = 2e$. Making use of the value of W_n given by relation (18.15), we have for the spectral series relation

$$\begin{aligned}\bar{\nu} &= \frac{2\pi^2\mu e^2E^2}{h^3c}\left(\frac{1}{k^2} - \frac{1}{n^2}\right)\\ &= \frac{2\pi^2 me^2 4e^2}{h^3c}\cdot\frac{1}{1+\dfrac{m}{M_{He}}}\left(\frac{1}{k^2} - \frac{1}{n^2}\right)\\ &= 4N_{He}\left(\frac{1}{k^2} - \frac{1}{n^2}\right).\end{aligned} \tag{18.20}$$

A particular series for ionized helium is then given by $k = 4$, $n = 5, 6, 7, 8, 9, 10, \ldots$. If we consider only the members of this series which have even values of n equal to $2n'$, where $n' = 3, 4, 5, 6, \ldots$, we can then write

$$\bar{\nu} = 4N_{He}\left(\frac{1}{4^2} - \frac{1}{(2n')^2}\right)$$

$$= N_{He}\left(\frac{1}{2^2} - \frac{1}{n'^2}\right), \tag{18.21}$$

a relation identical with that for the Balmer series except for the slight difference between N_H and N_{He}. In other words, Bohr's theory *predicts* that in the spectrum of helium, there should be lines almost identical with the Balmer hydrogen lines. Now, as a matter of fact, when helium is subjected to a strong electric discharge, spectral lines with wave-lengths almost equal to these hydrogen lines *are* observed.

When $k = 3$, we have another spectral series for ionized helium. On giving n successive integral values, wave-lengths are obtained which agree with lines, some observed by Fowler (England), others predicted by Rydberg (Germany), and at one time ascribed by both these experimenters to luminous hydrogen. Bohr's theory then leaves little doubt that the lines Fowler observed are due to helium.

Relation (18.20) also indicates that, since $E = 2e$, the Rydberg constant has the value $4N$ for ionized helium, or to be more exact, $4N_{He}$. Now it is a very significant fact that before the advent of the Bohr theory, Fowler had shown that, to express correctly the series relations for many enhanced spark lines (Chapter VII), the constant $4N$ was necessary. Bohr's view provided the explanation. Such lines are due to an element in the ionized state, while arc lines, as briefly indicated in Chapter VII, arise from the atom in the neutral state. Later spectroscopic investigation has shown that for the same element, there can exist *several* series of spark lines, for which the constants are 2^2N, 3^2N, $4^2N \ldots$. These exist in the case of atoms from which one, or two, or three or even more electrons can be removed. Singly, doubly, trebly, multiply

— ionized atoms can then exist, with corresponding constants $2^2N, 3^2N \ldots$ . .

It is evident, then, that Bohr's theory, which accepts quanta, which assumes that accelerated charges need not radiate, which quantizes angular momentum, is responsible for an amazing advance in the interpretation of facts in spectroscopy. Its successes give overwhelming support to the view that, for the satisfactory explanation of emission processes, a quantum theory is necessary. Bohr's theory provided the foundation on which later a vast structure has been erected by theoretical physicists. The Zeeman effect, the Stark effect, and the fine structure of spectral lines have been " explained " utilizing principles of a quantum mechanics. In many respects points of view have constantly changed, and are still changing, but throughout all the development so far as the origin of spectra is concerned, there remains at least one big fundamental conception, and that is, *the identification of spectral terms with energy levels.* While mathematical considerations cannot so readily be applied to the more complicated spectra of elements with several electrons, there can be no doubt that the difference in terms which gives the wave-number of a spectral line represents a difference in the energies of two states of the atomic system. In the next chapter this point of view will be further discussed, together with additional evidence in support of it.

168A. A Vector Model of the Atom and Further Interpretation of Spectral Series. — It will be recalled that according to Bohr's simple theory for the hydrogen atom, possible orbital paths were fixed according to the quantum condition

$$\text{angular momentum} = n \cdot \frac{h}{2\pi},$$

where n has integral values 1, 2, 3, ... Moreover, for any given orbit, the corresponding total energy, as shown by relation (18.12) depends on the value of n. We may, therefore, conveniently designate different terms or energy levels by values of n; or, alternately, we may describe the orbit in which an electron is

revolving by the value of n. As we have seen, such a picture enables us to interpret the various spectral series of hydrogen.

When we try to interpret the different types of series in an element like sodium, or to show why hydrogen spectral lines are in reality close doublets (see again, Fig. 4 (*a*), Plate VII), the details of this picture must be altered. This was first done by Sommerfeld (Germany) and W. Wilson (Great Britain) by a consideration of the elliptical orbits in which an electron may revolve about a nucleus. Mathematical analysis shows that, for *each* circular orbit, there is a group of elliptical orbits, *for all of which, to a first approximation*, the total energy is the same as that given by Bohr's simple theory for the circular orbit. To distinguish between the various elliptical orbits in a particular group of this sort, as well as between different groups, two quantum numbers are necessary. A single group is designated by n, the same number used to designate a Bohr circular orbit, but now n is called the *total quantum number*. The individual orbits in a single group are distinguished by the value of a number represented by k and called the *azimuthal quantum number*. This quantum number, which may have integral values ranging from 1 to n, is defined by the relation

$$\text{angular momentum} = k \cdot \frac{h}{2\pi}.$$

The mathematical analysis shows that

$$\frac{k}{n} = \frac{\text{minor axis of ellipse}}{\text{major axis of ellipse}}.$$

We see, then, that (1) in any given group, the most elliptical orbit is that for which $k = 1$, and (2) when $k = n$, we have the Bohr circular orbit with the same quantum condition as given on page 377. To describe a particular orbit, therefore, values of both n and k must be known, and we may conveniently designate the orbit by the symbol n_k. In Table C all possible orbits are given for $n = 1, 2, 3$ and 4.

TABLE C

Elliptical and Circular Orbits

n	k	orbit
1	1	1_1
2	1, 2	2_1, 2_2
3	1, 2, 3	3_1, 3_2, 3_3
4	1, 2, 3, 4	4_1, 4_2, 4_3, 4_4

A spectral line is emitted when an electron passes from any orbit to another one of lesser energy. In the case of hydrogen, each of the transitions from either 3_1, 3_2, or 3_3 to 2_1 or 2_2 gives rise to the same line (H_α) because (but again *only to a first approximation*) the total energy depends only on $n = 3$ for the first state and $n = 2$ for the second.

In the case of an element like sodium, however, the energy value depends on both n and k. In this element, of atomic number 11, radiation is emitted as a result of the transitions from one orbit to another of a single outer valency electron which revolves about the nucleus of the atom and 10 inner electrons. The combination of the nucleus, with its positive charge of 11 units, plus the 10 inner electrons, is sometimes called the *core* of the atom. If an electron revolves in a path which is always well outside the core, then the situation is very similar to that of the hydrogen atom, for the *outside* effect of the core on the valency electron is almost exactly the same as that of a single positive charge (the resultant of $+11$ on the nucleus and -10 on the inner electrons). If, however, the valency electron in its orbit penetrates the core, the case is very different from that of the hydrogen atom and the energy values are greatly altered as compared with those of hydrogen terms.

The chance of the valency electron penetrating the core depends on two things: (1) the value of n; (2) the degree of ellipticity of the orbit, that is, the value of k. If n is great enough, that is, if the orbits are far enough out, the electron will not penetrate the

core for any ellipticity. For smaller orbits, however, that is, for small values of n, the more elliptical the orbit, the greater the degree of penetration. Hence, the smaller k, the greater the penetration and the greater the departure of term values from corresponding hydrogen terms. Now the empirical analysis of such spectral series as those of sodium shows that S terms differ from hydrogen values more than P terms, and P term values more than D terms. The point is illustrated by the actual term values

TABLE D

Comparison of Some Hydrogen and Sodium Term Values

n	Hydrogen *	Sodium		
		S	P	D
3	12186	41449	24476	12276
4	6855	15709	11176	6900
5	4387	8248	6406	4412
6	3047	5077	4151	3062
7	2238	3437	2907	2248

given in Table D. It is natural, therefore, to associate S terms with values of $k = 1$; † P terms with $k = 2$; D terms with $k = 3$; and so on for other series of terms, which, for the sake of simplicity, have been omitted. For the same reason all n_1 electron orbits are often described by the symbol ns; n_2 orbits by the symbol np; n_3 orbits by nd. That is, capital letters are used to designate *term* values and the corresponding small letters to represent electron *orbits*.

First Selection Principle. — Although the wave-number of every spectral line is equal to the difference between two terms, the converse is not true. The difference between every two terms

* These are found at once from $\frac{N}{3^2}, \frac{N}{4^2}, \frac{N}{5^2}$, etc., where $N = 109678$.

† Subsequent theoretical developments have made it desirable to replace k by l, where $l = k - 1$. Thus, for S terms $l = 0$; for P terms, $l = 1$; for D terms, $l = 2$.

does not always give rise to a spectral line. For example, we do not encounter observed spectrum lines corresponding to differences between S and D terms. An examination of the lines which actually are observed leads to a rule, known as the first *selection principle*, according to which electron transitions take place only when the change in k (or in $l = k - 1$) is ± 1. For example, we observe spectral lines for transitions from $S(l = 0)$ to $P(l = 1)$ terms, or from $D(l = 2)$ to P terms but not from S to S or S to D. Occasionally, in the presence of strong electric fields, faint lines occur in violation of this principle.

168B. Doublet Fine Structure. — In section 73A it was pointed out that in the S, P and D spectra series of the alkali elements, the lines occur in pairs or doublets. In the energy level or term diagrams for these elements all P and D levels * are accordingly double, that is, for each value of n there are two sub-levels. In the short hand notation for spectral terms described in that section these sub-levels were distinguished by the use of subscripts. For example, the two sub-levels of the nth P term were there described as $nP_{\frac{3}{2}}$ and $nP_{\frac{1}{2}}$.

To interpret these sub-levels in terms of a model atom it is necessary to introduce still another quantum number. This is done (1) by ascribing to the valency electron a spin with an angular momentum equal to $\pm \frac{1}{2} \cdot \frac{h}{2\pi}$ (+ for rotation in one sense, − for rotation in the opposite); (2) by ascribing to the atom as a whole an angular momentum equal to $j \cdot \frac{h}{2\pi}$, where j is called the *inner quantum number*. To obtain the angular momentum of the whole atom, that is, the value of j, the spin momentum of the electron must be combined vectorially with the orbital angular momentum, which according to modern quantum mechanics is $l \cdot \frac{h}{2\pi}$. As already noted, l is related to the azimuthal quantum number k, by the simple relation $l = k - 1$. (Since a spinning electric charge

* In some of the elements the terms are too close to be separated.

gives rise to a magnetic field and an orbital motion of a charge is equivalent to a current with a resulting magnetic field, the spin and orbital momenta combine because of the interaction of these two magnetic fields.)

In the case of a single valency electron, because of the two possible directions of spin, for each orbit we obtain in general two values of j given by

$$j = l + \tfrac{1}{2} \quad \text{and} \quad j = l - \tfrac{1}{2}.$$

For S term values, for which $k = 1$ or $l = 0$, j can have only the value $\frac{1}{2}$, a negative value being inadmissible, hence S terms are single. This corresponds to the $^2S_{\frac{1}{2}}$ designation of section 73A. For all P terms ($k = 2$ or $l = 1$) j has values $1 + \frac{1}{2}$ and $1 - \frac{1}{2}$, that is, $\frac{3}{2}$ and $\frac{1}{2}$ and P terms are double. These correspond to the $^2P_{\frac{3}{2}}$ and $^2P_{\frac{1}{2}}$ designation. For D terms ($k = 3$ or $l = 2$) we have $^2D_{\frac{5}{2}}$ and $^2D_{\frac{3}{2}}$.

Second Selection Principle. — Transitions between different sub-levels are restricted according to this principle which states that only those occur for which the change in j is ± 1 or 0. For example, transitions $^2D_{\frac{5}{2}}$ to $^2P_{\frac{3}{2}}$; $^2D_{\frac{3}{2}}$ to $^2P_{\frac{3}{2}}$; $^2D_{\frac{3}{2}}$ to $^2P_{\frac{1}{2}}$ are allowable but not $^2D_{\frac{5}{2}}$ to $^2P_{\frac{1}{2}}$, for in the last case the change in j is $\frac{5}{2} - \frac{1}{2}$ or 2. Reference has already been made to this principle in connection with the discussion of a cadmium triplet on page 174.

The two selection principles may conveniently be summarized as follows.

$$\triangle k = \pm 1 \quad \text{and}$$
$$\triangle j = \pm 1 \quad \text{and} \quad 0,$$

where $\triangle k =$ change in k, and $\triangle j =$ change in j.

Doublet or Fine Structure of Hydrogen. — In section 90 it has already been pointed out that the analysis of the Balmer series lines of hydrogen (for example, either H_α^1 or H_α^2) by a Fabry-Perot étalon shows that each line is a close doublet (see again, Plate VII, Fig. 4 (*a*) and (*b*)). The doublet nature of these lines was first explained by Sommerfeld in connection with his work on elliptical orbits. When allowance is made for the fact that the mass of an electron varies with its velocity, it is not difficult to show that

the total energy for all elliptical orbits of the same total quantum number n is constant only to a first approximation. The solution which takes into account this relativity change of mass with velocity introduces a small correcting term whose value depends on k. The total energy of a 3_1 orbit then differs slightly from a 3_2 or a 3_3 orbit and the value of H_α^1 as obtained from the transition 3_3 to 2_2 is not quite the same as its value for the transition 3_1 to 2_2 or for 3_2 to 2_1. When the first selection principle ($\triangle k = \pm 1$) is applied to all possible transitions between $n = 3$ and $n = 2$ orbits, it is easy to see that H_α^1 should have three components, corresponding to changes in k values 3 to 2, 2 to 1 and 1 to 2. As the separation between two of these components is too small to be experimentally detected, only two are observed in actual practice — hence the doublet actually observed.

With the advent of the spinning electron and the more recent quantum mechanics, another and more accurate explanation of the fine structure of the hydrogen lines can be given. In this work both the relativity change in mass and the spin of the electron are taken into consideration. Strangely enough the two theories lead to the same numerical values of the energy levels. There is this difference, however, that more components are predicted by the newer theory. In the case of the fine structure of ionized helium (a hydrogen-like atom) the observed fine structure agrees better with the predictions of the second theory.

168C. Triplet and Associated Singlet System. — An outline only is possible of the interpretation of a triplet system such as that of cadmium (section 73A). The same quantum numbers are used, although not quite so simply, for with such elements we have to deal with two valency electrons. The resultant angular momentum of the atom as a whole is obtained: (1) by combining the spin momentum of each electron to get a resultant spin $S^* \frac{h}{2\pi}$; (2) by combining the two orbital momenta to obtain a resultant $L \cdot \frac{h}{2\pi}$.

* This letter S has nothing to do with the S used to designate a sharp term.

Since the spin momentum of a single electron is $\pm \frac{1}{2} \cdot \frac{h}{2\pi}$, the values of S can be only 1 or 0, that is, $\frac{1}{2} + \frac{1}{2}$ or $\frac{1}{2} - \frac{1}{2}$.

If the l values of each electron are designated l_1 and l_2, then the vectorial combination gives results such as are shown in Table E.

TABLE E

l_1	l_2	L *
1	1	0, 1, 2
1	2	1, 2, 3
2	2	0, 1, 2, 3, 4

To find the resultant angular momentum of the atom as a whole, that is, the value of $J \cdot \frac{h}{2\pi}$, the S values (resultant spin) must be combined vectorially with the L values (resultant orbital). If we confine our attention to L values equal to 0, 1 and 2, possible J values are given in Table F.*

TABLE F

S	L	J	Term
1	0	0	$^3S_0$
1	1	0	$^3P_0$
		1	$^3P_1$
		2	$^3P_2$
1	2	1	$^3D_1$
		2	$^3D_2$
		3	$^3D_3$
0	0	0	$^1S_0$
0	1	1	$^1P_1$
0	2	2	$^1D_2$

* In the vectorial combination of two integral quantities P and Q, where $P \geqq Q$, and the resultant can take only integral values, these values are $P + Q, P + Q - 1, \ldots P - Q$.

Just as in the case of a single valency electron values of $l = 0, 1, 2$ correspond to S, P and D terms respectively, so for two valency electrons $L = 0, 1, 2$, corresponds to S, P and D terms. It will be noticed also that for S (resultant spin) values equal to 1, we have a triplet system, that is, three sub-levels for P and D terms, whereas for S values equal to 0, we have a singlet system.

To pursue this subject further would bring us far beyond the scope of this book but it is hoped that the above outline will give the student a general idea of the interpretation of series spectra in terms of a model atom and quantum ideas.

168D. Hyperfine Structure (H.F.S.). — The separation of the components of a doublet may be so small that they appear as a single line in an ordinary spectrograph (for example, the lithium red line 6708 A which is in reality two wave-lengths $1^2S_{\frac{1}{2}} - 1^2P_{\frac{1}{2}}$ and $1^2S_{\frac{1}{2}} - 1^2P_{\frac{3}{2}}$); or they may appear as two widely separated lines (for example, the corresponding lines of rubidium with wave-lengths 8521 A and 8943 A). The term *fine structure* refers to this general complexity which can be explained in terms of the three quantum numbers, n, l and j (or n, L and J).

When many spectral lines are examined with a Fabry-Perot étalon or any instrument of sufficiently large resolving power and dispersion, they exhibit structure which cannot be explained in terms of these quantum numbers. A good illustration is provided in Fig. 3 (*b*), Plate VII, where the three components of the bismuth II line 6600 A exhibit *hyperfine structure*. Two causes may bring about H.F.S., both of them relating to the nucleus of the atom. The first cause operates when we have an element with two or more isotopes, that is, one made up of two or more different kinds of atoms. We have already called attention to the existence of the two H_α lines, separated by 1.79 angstroms, which are due to the two isotopes of hydrogen. In this case the separation is unusually large because of the large ratio of the masses (2 to 1). When the isotopic masses differ as little as they do for the element mercury (which has isotopes of atomic weight 196, 198, 199, 200,

202, 204) the component lines may be fairly numerous and close together.

The second cause * of H.F.S. arises from a magnetic moment, and so a quantized angular momentum equal to $I\frac{h}{2\pi}$, associated with the nucleus of the atom. By a still further extension of the combination of atomic vectors, I is combined with J values to give quantized hyperfine energy levels, distinguished by different values of a resultant F. The establishment of such levels in agreement with observed hyperfine structure is one of the fields of research in modern physics. Its importance lies in the information it gives about the nucleus of the atom.

* Some authors (Prof. H. E. White, for example, in his excellent book an *Introduction to Atomic Spectra*) refer to this as the true cause of H.F.S. and call the other, isotope structure.

CHAPTER XIX

RADIATION POTENTIALS; ABSORPTION AND BAND SPECTRA

169. Radiating Potentials. — In the case of the hydrogen spectrum, we have seen in the last chapter that, according to Bohr's theoretical considerations, the empirical spectral series relation

$$\bar{\nu} = \frac{N}{k^2} - \frac{N}{n^2}$$

is an expression of the frequency condition,

$$h\nu = W_n - W_k$$

On this view, in support of which much evidence was given, a spectral term is therefore a measure of the energy of a corresponding state or level of the atom. When we turn to the series relations of other elements, *for all of which wave-numbers are expressed as the difference between two terms*, it is natural to assume that the same general ideas hold, that is, that the spectral series relation is just a statement of the frequency condition. Each term then corresponds to a particular energy state or level, the highest numerical term being a measure of the energy of the atom in its normal state. (see page 378). These energy changes are the result of the displacement of an outer *valency* electron (and possibly, in special cases, more than one) from its normal position, because of some external exciting agency. In sodium, for example, there is good evidence that the nucleus is surrounded by eleven electrons, changes in the level of the outermost one being responsible for the emission of observed spectral lines. If the atom is only feebly excited, this outer electron may be displaced just enough to cause an energy change from the most stable to the next level, in which case the subsequent radiation, when the electron reverts to its normal position, consists of a single wave-length, the familiar yellow

D line. (Actually two lines D_1 and D_2 close together are emitted, thus showing the existence of two sub-levels close together.)

The soundness of the view which identifies terms with energy levels may be seen by a consideration of the following numerical example. The D lines of sodium are the first members of a principal doublet series described by the relation

$$\bar{\nu} = 41449 - \text{variable term.}$$

For the D_1 line, $\lambda = 5896$ A, we have

$$\bar{\nu} = 41449 - 24493.$$

For the D_2 line, $\lambda = 5890$ A, we have

$$\bar{\nu} = 41449 - 24476.$$

If, now, we take either of these wave-lengths, and re-write this relation in energy units, we have (using D_1)

$$h\nu = hc\bar{\nu} = hc(41449 - 24493).$$

Suppose, now, that this amount of energy is communicated to a sodium atom (which subsequently radiates 5896) as a result of bombardment by an electron. The potential V through which the electron must fall to acquire the necessary amount of energy is then given by

$$Ve = hc\bar{\nu} = hc(41449 - 24493),$$

from which, if we substitute the values of e, h and c already given we obtain

$$V = 2.093 \text{ volts.}$$

If the term values for D_2 are used, we obtain $V = 2.095$ volts. Before sodium vapour can be excited to emit the characteristic yellow lines, therefore, each atom must have communicated to it an amount of energy equal to that possessed by an electron which has fallen through a potential difference of this amount. But this is something which may be tested by actual experiment, using a method similar to that described, in (f) of section 68, for magnesium. Results obtained give the value of $V = 2.12$ volts, in excellent agreement with the calculated value.

The case of the D lines of sodium is, in reality, just another illustration similar to that given in the last chapter when the emission of a single line spectrum of mercury vapour was interpreted on a quantum basis. The additional point which is now being emphasized is that the quantum is emitted as a result of a change in the energy of an atom from one value or level, to another, brought about by an external agency, radiation resulting when the system reverts to the original level.

The value of the energy difference corresponding to the emission of a spectral line is frequently expressed in terms of the potential difference through which an electron must fall in order to acquire such an amount of energy, and hence, as so many volts. In the example worked out, the emission of D light by sodium is the result of an energy change equivalent to 2.1 volts, or corresponds to what is called a *radiating* or *excitation potential* of 2.1 volts.

The second member in the series for which D_1 is the first is λ 3303, described by

$$\bar{\nu} = 41499 - 11181.$$

By applying exactly the same method, the corresponding radiating potential is found to be 3.74 volts. Actually, as already pointed out more than once, it has been shown experimentally that more and more lines appear with increasing energy of colliding electrons, and there are numerous cases showing the agreement between experimental and calculated values.

170. Ionisation Potential. — In any spectral series the values of the *variable* term gradually decrease, eventually becoming equal to zero. The limit of the series obviously represents the lowest energy level to which, for that particular series, the atom always reverts. When there are several different series for a given atom, as there invariably are, then the largest numerical limit must correspond to the normal state of the atom, and hence must be a measure of the work necessary to remove the electron to the zero level, that is, from its normal state entirely away from the attraction of the nucleus. When the electron has been so removed, the

atom is ionized, and the work necessary to do so, expressed in equivalent volts, is called the *ionisation potential.* Obviously its value may be calculated in the same manner as has been done for radiating potentials. Thus, to consider sodium once more, of the various spectral series for this element, the limit of the principal series has the largest numerical value, namely, 41449. Hence if V is the ionisation potential,

$$Ve = h.c.\ 41449,^*$$

from which $$V = 5.12 \text{ volts.}$$

Now there are various ways of observing experimentally the value of V for which colliding electrons ionize a gas or vapour. In the case of sodium, the value observed is 5.13 volts, showing again remarkable agreement between theory and experiment.

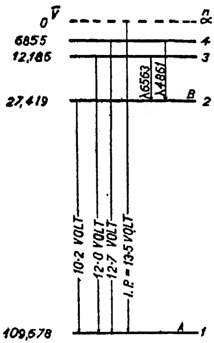


FIG. 215.

171. Energy Diagrams. — It should now be evident that the special series diagrams, such as those shown in Figs. 94 and 96, may also be called energy diagrams. Indeed, that is the usual name given to such figures, on which, frequently, excitation potentials are marked. Thus Fig. 215, which is identical with a portion of Fig. 94, is an energy diagram for the hydrogen atom showing the wave-numbers of a few energy levels, excitation potentials, and wave-lengths. Fig. 216 shows only a very few of the numerous levels for the sodium atom with the same information as in the case of hydrogen.

172. X-Ray Spectra. — In passing, it may be noted that analysis of X-ray spectra shows that in that field series also exist, and

* Since the value of $\frac{hc}{e} = \frac{1}{8102}$, this and the corresponding similar relation for radiating potentials always reduce to

$$\frac{\text{wave-number}}{8102} \quad \text{or} \quad \frac{12344}{\text{wave-length in angstroms}}$$

that individual wave-lengths may be expressed in terms of differences in energy levels. As far as origin is concerned there is much evidence to support the view that such X-ray energy changes arise from the removal of an inner electron close to the nucleus to a position farther away, while radiation results when, in general, another electron "drops" back to take the place of the displaced one. In the case of optical spectra it will be recalled that an atom is excited by the removal of an outer valency electron to one of many possible orbits, so-called *virtual* orbits, radiation resulting when the electron returns to an orbit nearer to the nucleus. It is significant that in X-ray spectra there are *progressive* changes as we pass across the periodic table of elements, whereas in optical spectra there are periodic changes. This is in line with the general theory of the origin of these two classes of radiation, for, in chemical combinations, it is the outer valency electrons which play an important part, and of course, chemical properties of elements vary markedly as we pass across the table.

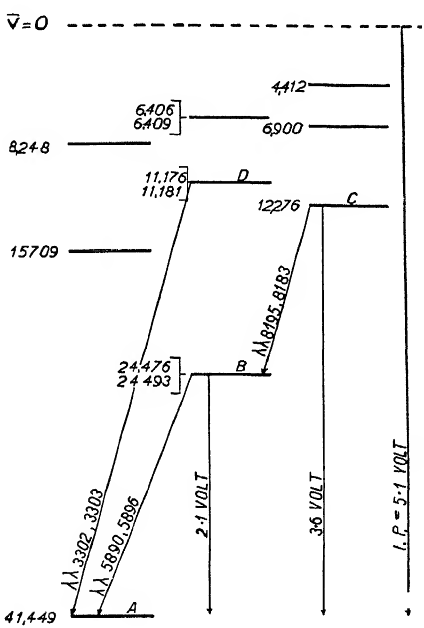


FIG. 216.

On the other hand, on this view it is to be expected that there should also be marked similarities between optical and X-ray spectra, especially in the case of lighter elements where the total number of electrons is comparatively small. That such is actually the case has been established, especially by the work of Millikan

and Bowen, who have shown that certain quantitative laws governing the separation of X-ray doublets may be extended to include the case of optical doublets for many of the lighter elements.

173. Resonance Radiation. — Some years before the advent of Bohr's theory it was known that certain vapours, when illuminated by light of a suitable kind, re-emitted light of the same kind in all directions. To give a concrete illustration, Wood of Johns Hopkins University, showed that, when sodium vapour was illuminated by incident D light, the vapour re-emitted D light in all directions. To such a re-emission he gave the name *resonance radiation*. On Bohr's theory the explanation is not far to seek, for it is evidently just a case of supplying the energy necessary to raise an electron from the normal to the next level, not by colliding electrons but by incident radiant energy. A quantum $h\nu$ is absorbed, (where ν is the frequency of the incident light), the energy is just sufficient to change the level from the innermost to the next, after which the reverse change results in the emission of a quantum of the same frequency. In itself this fact is not necessarily a confirmation of the truth of Bohr's theory, for, on the older views, it may be considered (as has already been done in Chapter VII) just a case of the well-known phenomenon of resonance. When, however, one considers the fact that excitation of sodium vapour by light of wave-lengths λ 3303 and λ 3302 (second members of the principal sodium series), also causes the emission of the D lines, and when one understands certain facts relating to absorption spectra, it will be seen that the idea of energy levels is strongly supported.

174. Absorption Spectra. — To understand this additional evidence, suppose a beam of light from a source emitting a wide continuous range of wave-lengths traverses a column of sodium vapour in an exhausted tube. When such a beam, on emergence from the tube, is analyzed by a spectroscope or spectrograph, the D lines and other members of the principal series appear as narrow black absorption lines, whereas there is *no* absorption of any of the members of the sharp and the diffuse series. The explanation of

the absorption of one series and non-absorption of others can readily be given in terms of energy levels. Since in normal, that is, unexcited sodium atoms, the valency electron is in an orbit corresponding to the innermost level (marked A in Fig. 216), it can be displaced as a result of absorption of light, only by those wave-lengths which correspond to a transition from some higher level to this normal level. A glance at the energy diagram will show that this is the case for all members of the principal series, such as 5896 3303, but is not so for lines of the other two important series. Thus wave-lengths 8195 and 8183, the first pair of the diffuse series, arise from a transition from level C* (Fig. 216) to level B.* Hence the absorption of these wave-lengths could only raise the level from B to C, and, therefore, since in *normal* sodium, no level B exists, such absorption is not possible. On the other hand, any of the members of the principal series can by absorption raise the level from A to some higher level and are, therefore, present in the absorption spectrum. For the same reason, absorption by such wave-lengths as 3303 and 3302 can put the atom in such a state that it may subsequently radiate the D lines. Thus the absorption of 3303 and 3302 raises the level to D, from which the electron may return to A in stages, the last one being that from level B to level A, which corresponds to the emission of D light.

A striking confirmation of the same ideas is provided by experiments on the absorption spectrum of hydrogen. Normal unexcited hydrogen does not absorb the members of the Balmer series, because, before any of such wave-lengths can be absorbed, the electron must be in an orbit corresponding to energy level B, Fig. 215. In normal atoms, however, the electrons are in orbits corresponding to level A. On the other hand, if, during the absorption experiments, hydrogen is electrically excited (for example, by passing a discharge through it), it was shown by Ladenburg that the Balmer series lines are then absorbed. In this case, as a result of electrical excitation, a large number of hydrogen atoms are in a state where the electron has been removed from the nor-

* In reality, we have two levels close together at B and at D. To keep the diagram as simple as possible only one has been drawn.

mal orbit to that corresponding to level B, and absorption of the members of the Balmer series is then possible. It is interesting to note that this experiment was first performed before the idea of energy levels came into existence.

175. Band Spectra. — In Chapter VII, where a brief reference was made to band spectra, it was pointed out that a single band is characterized by a group of lines so close together that in an instrument of small dispersion the lines fuse together, a whole group of such bands then having a fluted appearance, as shown in Fig. 3, Plate II. With sufficient dispersion, the individual members of a single band are apparent, frequently separated by intervals which become less and less as an edge, called the band head, is approached. Examine, again, the individual band shown in Fig. 4, Plate II. Before the days of quantum theories, a certain amount of analysis of band spectra had been undertaken, notably by Deslandres (France) and it had been shown, among other things, that the wave-numbers for band heads could be expressed by an empirical relation involving constants and two variable integral numbers. No adequate explanation of the origin of bands was forthcoming, however, until the application of quantum principles was made. The problem, in some respects, is more complex than that of line spectra, for in all cases band spectra result from a luminous substance in the *molecular* state. Remarkable progress has been made, however, by considering radiation as the result of changes in energy values for different states of the molecule, and by quantizing quantities, just as in the case of the atom and line spectra. In the remainder of this chapter a brief summary of the theoretical interpretation will be given.

Consider a diatomic molecule, that is, a union of two atoms, the nuclei and surrounding electrons being so arranged that equilibrium and stability results. An external "exciting" agency may bring about three general changes in such a system: (1) one or more electrons may be displaced from a lower to a higher energy level as in the case of an atom, that is, there may be an electronic change; (2) the two nuclei may be displaced so as subsequently to vibrate

with certain amplitudes, that is, with definite amounts of *vibrational* energy; (3) the system may be made to rotate about an axis of symmetry with increased *rotational* energy. A molecule may be excited, therefore, and subsequently radiate because of changes in any or all of these three kinds of energy. In the most general case, when a quantum $h\nu$ is radiated, we may write

$$h\nu = h\nu_e + h\nu_v + h\nu_r, \tag{19.01}$$

when $h\nu_e$ is the energy arising from an electronic change in level, $h\nu_v$ from a vibrational change, $h\nu_r$ from a rotational change.

We shall examine briefly three different kinds of bands.

(1) **Rotation Band.** — In quantizing vibrational and rotational changes, the same general ideas as employed for line spectra are utilized. Thus, when considering the rotation of the two atoms of a diatomic molecule about an axis perpendicular to the line joining them, the angular momentum is restricted to integral multiples of $\frac{h}{2\pi}$. Hence, if

I = moment of inertia of the system about the axis of rotation,

ω = angular velocity,

$I\omega$ = angular momentum $= m \cdot \frac{h}{2\pi}$, where m has integral values.

From this relation, we may write for W_m, the energy of the mth state,

$$W_m = \tfrac{1}{2}I\omega^2 = \frac{m^2h^2}{8\pi^2 I}. \tag{19.02}$$

According to modern quantum mechanics m^2 is replaced by $J(J+1)$, where J has integral values 0, 1, 2, 3, etc. We have, therefore, a series of quantized rotational levels defined by this quantum number J. In what is called a *pure rotation band*, radiation (or absorption) takes place solely because of transitions from one of these levels to another. Since in such transitions J changes by 1 only, $\bar{\nu}$, the wave-number of each of the emitted (or absorbed) wave-lengths constituting the band is given by

$$\bar{\nu} = \frac{1}{\lambda} = \frac{\text{energy change}}{hc}$$

$$= \frac{1}{hc}\left[(J+1)(J+2)\frac{h^2}{8\pi^2 I} - J(J+1)\frac{h^2}{8\pi^2 I}\right]$$

$$= \frac{2h}{8\pi^2 Ic}(J+1)$$

$$= 2B(J+1), \text{ where } J = 0, 1, 2 \ldots, \tag{19.03}$$

where $$B = \frac{h}{8\pi^2 Ic} = \frac{27.7 \times 10^{-40}}{I}, \tag{19.04}$$

since $h = 6.55 \times 10^{-27}$ and $c = 2.998 \times 10^{10}$.

In the case of the hydrogen halides *HF*, *HCl*, *HBr*, and *HI*, bands conforming closely to this law have been observed in absorption in the far infra-red region, in a region of wave-lengths of the order of 0.1 mm. Their presence in the far infra-red where frequencies are so much smaller than in the visible (compare, for example the frequency for $\lambda = 0.01$ cm. with that for $\lambda = 5 \times 10^{-5}$ cm.) tells us that rotational energy levels are much closer than those we encounter in line spectra, that is, than electronic levels.

For *HCl* the *empirical* formula giving the wave-number of each line of the observed rotation band is

$$\bar{\nu} = 20.793(J+1) - 0.00163(J+1)^3.$$

If we neglect the small second term on the right hand side, this expression is of exactly the same form as relation (19.03), $2B$ being equal to 20.793. As a matter of fact, when allowance is made for slight changes in the distance apart of the two rotating nuclei, correcting terms must be added to the simple expression given in (19.03). In presenting the fundamental ideas of band spectra it has been thought wiser to omit such correcting terms, whose magnitudes, as already pointed out in the case of *HCl*, are small compared with $2B(J+1)$.

Fig. 217 gives the levels or term values for an observed rotation band, the first four members of the band being represented by vertical lines. The actual separations of successive levels are

(in wave-numbers) 20.79, 41.58, 62.33, and 83.07. Note (1) how small these are compared with the separation of levels encountered in line spectra, such as in Figs. 215, 216; (2) in conformity with relation (19.03), the increasing separation of the levels, in the ratio 1 : 2 : 3 : 4 etc.

FIG. 217.

(2) **Vibration-Rotation Bands.** — In addition to the rotation of a diatomic molecule about an axis perpendicular to the internuclear axis, the two nuclei may vibrate about equilibrium positions and so possess also vibrational energy. This energy is quantized in terms of the unit $h\nu_e$ where ν_e is the frequency for vibrations of very small amplitude. In this case the quantum number is represented by the symbol v and according to quantum mechanics, the vibrational energy for any state v, is given by the relation

$$\text{vibrational energy} = (v + \tfrac{1}{2})h\nu_e,^* \text{ where } v = 0, 1, 2 \ldots \quad (19.04)$$

To obtain the corresponding *term* values, the energy values as usual must be divided by hc. Thus we have a series of vibrational levels whose values to a first approximation, in wave-numbers, are (if we put $v = 0, 1, 2, 3, \ldots$)

$$\frac{1}{2}\frac{h\nu_e}{hc}, \quad \frac{3}{2}\frac{h\nu_e}{hc}, \quad \frac{5}{2}\frac{h\nu_e}{hc}$$

Associated with *each* of these is a series of rotational levels. In Fig. 217A, for example, we have represented 3 vibrational levels distinguished by $v = 0$, 1 and 2, and associated with each of these, 4 rotational levels, distinguished by the J values, 0, 1, 2, 3. A particular rotational level may, therefore, conveniently be referred to as a (v, J) term, where v can have values 0, 1, 2, . . .,

* Here, again, this is only to a first approximation and correcting terms must be added to take into consideration the fact that the vibrations are anharmonic, not harmonic. The more exact relation is

$$(v + \tfrac{1}{2})h\nu_e - x_e(v + \tfrac{1}{2})^2h\nu_e + y_e(v + \tfrac{1}{2})^3h\nu_e + \cdots,$$

where x_e and y_e are very small constants.

and J, 0, 1, 2, 3. Actual experiment shows the existence of wave-lengths arising not only from transitions between two successive rotational levels associated with the same vibrational level (that is, a pure rotation band) but also of wave-lengths corresponding to transitions from a rotational level of one vibrational group to a rotational level of a different vibrational group. In such transitions values of v may change by 1 or 2 or 3, but changes in J values

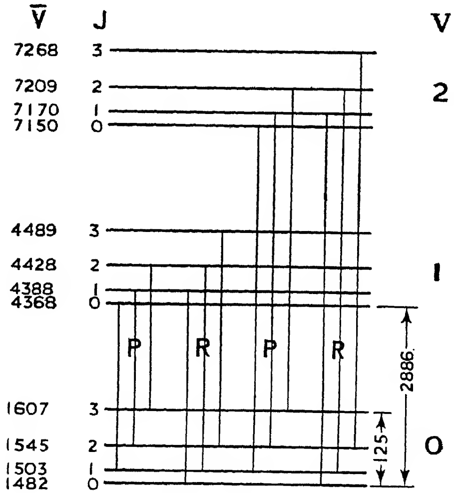


FIG. 217A.

are restricted to $\pm$ 1. When the transitions are between rotational levels associated with different vibrational levels, the bands are called *vibration-rotation*. All the wave-lengths or lines corresponding to the same vibrational change constitute a single vibrational band. Thus, in Fig. 217A, all the vertical lines joining different rotational levels of the $v = 1$ group to those of the $v = 0$ group, represent the wave-lengths which make up the (1, 0) band. Similarly, the (2, 0) band is made up of all the wave-lengths arising from transitions between rotational levels of the $v = 2$ group to those of the $v = 0$ group.

The wave-numbers marked in Fig. 217A are actual observed values for HCl^{35}.* The fact that the observed spectra, although still in the infra-red, are of much shorter wave-length than those for pure rotation, tells us that for vibration-rotation bands we are dealing with considerably larger energy values. In the case of HCl (Fig. 217A) this is at once apparent from the much greater wave-number separation of two vibrational levels as compared with the separation of rotational levels.

It will be noticed in Fig. 217A, that the lines representing the (1, 0) and the (2, 0) vibration bands are in two groups, marked P and R. The R group or *branch*, as it is usually called, is composed of all lines for which the J value of the lower vibrational level is 1 less than that of the upper level whereas the P branch is composed of those for which the value is 1 more than that of the upper. In more complicated spectra there is a third branch Q composed of lines for which the change in J is 0.

Isotope Effect. It has been stated above that the wave-numbers in Fig. 217A are those of actual levels for band spectra of HCl^{35}. It is interesting to note that, when a vibration-rotation band of HCl was first discovered in 1919 by Ames (United States) the appearance of each line in the band suggested a doublet. This was soon shown by Loomis (United States) and others to be due to two components arising from the two isotopes of chlorine, that is, one component for HCl^{35} and a weaker one for HCl^{37}. The student should see without difficulty that the magnitude of the vibrational frequency, as well as that of I, the moment of inertia (and hence of B, in relation 19.04) varies with the mass of each nucleus. Hence, given known masses, it is possible to predict the isotopic separation between the corresponding members of the two bands, or conversely, from the observed separation, to calculate the ratio of the masses. In the case of HCl^{35} and HCl^{37} good agreement was obtained between observed and predicted values.

* Chlorine has two isotopes of atomic weight 35 and 37, hence there is HCl^{35} and HCl^{37}.

After the discovery of heavy hydrogen or deuterium, no difficulty was experienced in obtaining H^2Cl^{35} and H^2Cl^{37}. It then became possible to use the isotopic separation of band spectra in order to compare the mass of the atom of deuterium (that is, of H^2 or D) with that of ordinary hydrogen (H^1). This was done by Hardy, Barker and Dennison (all of United States) by observing the isotopic separation in the case of H^1Cl^{35} and H^2Cl^{35}. A result in excellent agreement with other methods was obtained.

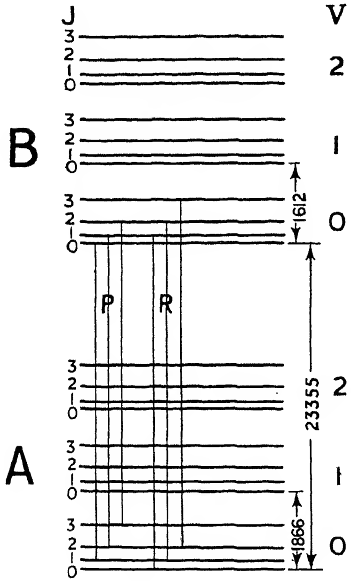


FIG. 217B.

(3) **Electronic Band.**—As already pointed out, the energy changes involved in vibration-rotation bands are so small that all the observed lines are in the infra-red region. To obtain band spectra in the visible or ultra-violet region much greater energy changes are necessary. Such changes arise when the energy of the molecule is altered by electronic displacements just as in the case of ordinary atomic line spectra. For all the levels shown in Fig. 217A the molecule is in the same electronic configuration or *all* the levels there represented correspond to the same electronic state. If increases in energy are brought about by electronic changes, the molecule may be in different quantized electronic states which we shall call A, B, C, etc.

For *each* of these states there is a series of vibrational states distinguished by v values = 0, 1, 2, 3, . . .; and for each vibrational state, a series of rotational levels or states distinguished by J values = 0, 1, 2, 3, ... In Fig. 217B we have represented two electronic states of a molecule, marked A and B; for each of

these, three associated vibrational states; and for each vibrational state, four rotational levels. As a rule the actual levels would be more numerous than those shown in this figure. Some idea of the relative separation of the various levels may be had from the wave-number differences (relating to certain *CuH* bands) marked on the diagram. Thus the energy of the $v = 0$ vibrational level of electronic state B differs from that of the $v = 0$ vibrational level of state A by about 23355 cm.$^{-1}$ (that is, wave-numbers). Compare this with the separation of the two vibrational levels, $v = 1$ and $v = 0$ of state A, which is only 1866 cm.$^{-1}$. The wave-number separation of the $J = 0$ and $J = 1$ rotational levels of the $v = 0$ vibrational state, for electronic state A, is not marked on the diagram but is actually about 16 cm.$^{-1}$.

When, therefore, transitions occur from levels of state A to those of state B the wave-number changes are so great that the emitted (or absorbed) lines are usually in the ordinary visible or ultra-violet region. (The wave-number difference 23355 corresponds to a wave-length of 4282 A.) In such transitions v may change by ± 1, ± 2, ± 3 . . ., J by ± 1 or 0.

A single band is composed of all lines corresponding to transitions from one vibrational level to another, the band usually being described by giving the value of v for the upper electronic state, side by side with that of v for the lower. Thus the band (2, 0) refers to the band arising from transitions from the $v = 2$ upper state to the $v = 0$ lower. In Fig. 217B lines of the P and R branches for the (0, 0) band are marked with P and R.

Because of the closeness of the rotational levels the individual lines in a band are frequently not separated, unless sufficiently high dispersion is used. With low dispersion a single band has generally the appearance of a sharp edge on one side, where the individual lines are crowded together most closely, with a gradual decrease in intensity (a degrading) to the other side, where sometimes individual lines may be seen. Fig. 3, Plate II, shows the appearance of a dozen or so unresolved bands. Under very high dispersion the individual lines composing the band are separated. A beautiful example of this (the 2430 band of magnesium hydride),

for which the author is indebted to Dr. R. W. B. Pearse, of the Imperial College, London, England, is given in Fig. 4, Plate II.

An electronic band *system* is a collection of all the individual bands arising from transitions from one electronic state to another. For example, in the band spectra of nitrogen, there is a *first positive system* arising from transitions from a state B to a state A; a *second positive system*, from state C to state B; and a *fourth positive system*, from state D to state B, as well as other systems.

Sequences and Progressions. — Sometimes in the simpler band systems groups are repeated in such a way as to suggest some regularity. On analysis it is found that each group consists of bands for which the numerical difference between the v of the upper electronic state (usually represented by v') and the v of the lower (represented by v'') is the same for each band. Such a group is called a *sequence*. For example, in one of the systems of bands due to CN, there are four sequences corresponding to values of $v' - v'' = +1, 0, -1$, and -2. In a single sequence, such as the -2, there are at least six separate bands corresponding to $v' = 0, 1, 2, 3, 4, 5$ and $v'' = 2, 3, 4, 5, 6$ and 7.

Sometimes bands are also grouped in *progressions*, of which there are two kinds: (1) a v' progression made up of all bands with the same v'' value but different v' values; (2) a v'' progression, in which the v' value is constant and v'' values are variable.

While only the bare elements of the analysis of band spectra has been given, it should be sufficient to enable the reader to understand the lines along which much progress in this field has been and is being made. From measurements of the wave-lengths involved, the constants occurring in the band spectra relations can be evaluated, and hence a determination made of such quantities as the distance between the nuclei of a diatomic molecule, or of the amount of work necessary to dissociate such a molecule. Moreover, as already indicated in the case of the two kinds of hydrogen, valuable information regarding isotopic masses can be obtained from an analysis of band spectra.

CHAPTER XX

THE DILEMMA

176. In the discussion of the relative merits of the corpuscular and the wave theory of light as given in Chapter I, it was pointed out that the discovery of interference not only demanded a wave theory of some kind, but also paved the way for a satisfactory explanation of the outstanding optical problems of the nineteenth century. Moreover, a large portion of this book has been devoted to a study of phenomena which find a ready interpretation on a wave theory, and indeed, some of which can only be explained in such a way. On the other hand, in the chapters immediately preceding this one, irrefutable evidence has been given that there are other optical phenomena, notably those referring to the interaction between radiation and matter, which can only be explained in terms of a quantum theory. Now in some respects, these two theories are contradictory. Consider, again, the outstanding differences.

177. According to the wave theory, radiant energy spreads out *continuously*, the amount received per sq. cm. on any given surface constantly decreasing, the farther the surface is removed from the source. On the other hand, on the quantum theory, radiant energy has an atomicity, the fundamental unit being of magnitude $h\nu$. Moreover, while certain facts can be explained on this theory by restricting the discontinuity to processes of emission and of absorption, the empirical laws of such a phenomenon as photo-electricity seem to require the propagation of light in quanta *which maintain their identity throughout their journey*. The Compton effect (section 179) points to the same conclusion. In such cases we have to do with *discontinuous* wave-fronts, and a flight of what are variously called light quanta (after Einstein), light corpuscles, or *photons*.

Our position, then, is that apparently certain facts in light can only be explained on a wave theory, certain others, only on a quantum theory, which in many respects is similar to the old corpuscular theory of Newton.

It will be recalled also that in applying quantum ideas, contradictions between old and new views were introduced in regard to dynamics. Classical electromagnetic theory, based on Newtonian dynamics, said that accelerated charges should radiate; quantum theory postulates that this need not be the case. As long as Planck applied to bodies of atomic dimensions Newtonian mechanics, the facts of black body radiation could not be deduced. When a departure from classical conceptions was made, however, a correct interpretation of this problem was given.

Now an introductory text in light is not the place to discuss dynamical principles in detail, but, in general, the student should have no difficulty in realizing that the application of Newtonian laws of mechanics to atoms and molecules does not always account for observed facts. And yet there is no doubt of the amazing accuracy with which such laws apply to large aggregations of molecules, that is, to those masses whose motions are ordinarily considered. (Consider, for example, the exactness with which predictions can be made regarding the motion of astronomical bodies.)

178. The modern physicist is then faced with a dilemma. There seem to be two contradictory theories, one necessary for the interpretation of one group of facts, the other for a second group. Can both be right? Fortunately, such a situation is by no means an unmixed evil, for a scientist usually grasps both horns of a dilemma in his endeavour to wrest from Nature secrets yet denied him. Apparent contradictions sometimes imply the lack of some larger outlook by means of which conflicting views can be shown to be but two different aspects of the same co-ordinating principle. The finding of such a principle in connection with the question under discussion is one of the big problems whose solution is being sought by the ablest minds of this generation; and indications are not wanting that the solution is forthcoming.

Much progress has been made along lines which will be indicated briefly in the next few pages. While mathematical physicists are largely responsible for the advances made, new ideas have constantly been related to and tested by experimental facts, with some very remarkable results. Outstanding in the recent work is the experimental evidence in support of the following significant and complementary statements: (1) light corpuscles or photons, in encounters with electrons, may behave like particles of matter; (2) material particles such as electrons, under certain conditions, act like a group of waves.

179. The Compton Effect. — The idea underlying the first statement was used by A. H. Compton, Professor of Physics at the University of Chicago, in connection with the interpretation of results obtained when X-rays are scattered by free electrons. Compton, following the pioneer work of J. A. Gray (Canada), showed that, when a beam of monochromatic X-rays is scattered on passage through certain substances, the scattered radiation is of slightly longer wave-length than the incident. Since a longer wave-length corresponds to a lower frequency, the energy unit $h\nu$ in the original beam must be greater than that of the scattered beam by an amount which can readily be found from observations of the wave-lengths concerned. To interpret his observations Compton assumed, (1) in a collision between a quantum and a free electron, energy is communicated to the electron, thus, by the law of conservation of energy, giving rise after the collision to a quantum $h\nu'$ of less energy; (2) a quantum possesses *momentum* of amount $\frac{h\nu}{c}$,* the law of conservation of momentum being

* When light is incident on a surface, a pressure p is exerted. In the case of a black body, experiment shows that the pressure is equal to the energy per unit volume in the incident beam. If, therefore, n quanta per second are incident on a surface of 1 sq. cm., $p = n \cdot \frac{h\nu}{c}$, since each quantum possesses $h\nu$ units of energy and since the velocity of the incident radiant energy is c. Now if particles collide against a surface, the pressure

applicable to the collision, just as in the case of colliding particles. When the energy and momentum equations are then written down in the usual way, it is easy to derive an expression for the new frequency ν' in terms of the original value ν, and of the direction in which scattering is observed. The experimental results were found to be in remarkable agreement with the deduced expression.

180. Electron Waves. — That under certain circumstances electrons behave as if they were a group of waves, has been shown in two different ways, to each of which a brief reference will be

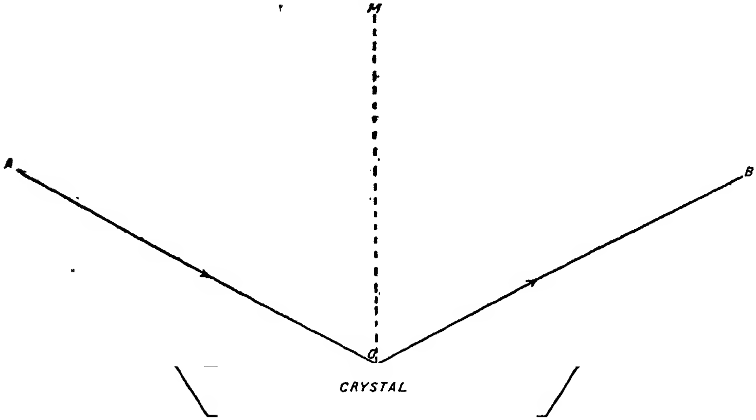


FIG. 218.

made. In the first method, used by Davisson and Germer, working in the Bell Telephone Laboratories, New York, a beam of electrons of a known speed was projected obliquely against the face of a nickel crystal and the intensity of the reflected beam was observed in various directions. Two outstanding results were ob-

exerted is equal to the rate of change of momentum. Hence, if we compare a quantum to a colliding particle, the change of momentum per second is equal to the pressure and therefore $= n \cdot \frac{h\nu}{c}$. In other words, each quantum may be thought of as having $\frac{h\nu}{c}$ units of momentum.

tained: (1) for electrons of a fixed velocity, the intensity of the reflected beam is a maximum in the direction OB, Fig. 218, where OB makes the same angle with the normal to the face of the crystal (represented by MO) as does the incident beam in the direction AO; (2) if the velocity of the incident beam is varied, as may readily be done by the method described in (f) of section 68, *the intensity of the reflected beam in the direction OB goes through maximum and minimum values. In other words, in the same direction, electrons of some speeds are strongly reflected, of others feebly reflected.*

Now when a beam of X-rays containing many wave-lengths is incident on the surface of a crystal in a manner for which Fig. 218 will serve as an illustration, certain wave-lengths only are reflected in the direction OB. Because of the path difference which exists between the X-ray disturbances reflected from successive parallel layers of atoms, we have to deal with the superposition of beams differing in phase, and hence the resultant reflected beam has a maximum intensity only for certain wave-lengths. Thus, if d is the distance between two such successive layers, and $\angle MOB = \theta$, there is a reflected X-ray beam of appreciable intensity *only* for wave-lengths which satisfy the relation

$$n\lambda = 2\,d\cos\theta,\text{* (where } n \text{ is integral).} \qquad (20.01)$$

If then d is known and θ is observed, the corresponding wave-lengths may readily be calculated.

Now, Davisson and Germer, making the assumption that electrons were waves, or were associated with a group of waves, reasoned that such a relation should hold for the reflection of electrons also. The fact that electrons of some velocities are strongly reflected, others not, follows at once *if the wave-lengths of such waves depend on the velocity of the electron.* On this hypothesis, then, the wave-length associated with a velocity which gave intense reflection can be calculated from the above relation (20.01). As an example of the results obtained by such calculations, these experimenters found for 54 volt electrons, $\lambda = 1.67$ A, for 65

* The path difference in question $= 2d\cos\theta$. Compare the expression for path difference in the case of a Fabry and Perot Étalon.

volts, $\lambda = 1.52$ A. If this were the whole story, no one could be blamed for being skeptical about the existence of electron waves. Fortunately, it is not, for, as a matter of fact, these results are a confirmation of results predicted by what is known as *wave mechanics*.

181. Wave Mechanics. — Wave mechanics is the name given to a new way of interpreting physical phenomena which gives promise of providing the fundamental principle that will reconcile the conflicting views we have been considering. To explain the connection between the experimental work of Davisson and Germer and this theory, a brief digression is necessary. In Chapter XVIII, dealing with the rise of the quantum theory, attention was directed to the initial successes of Bohr's theory of the origin of spectra and to his manner of quantizing the atom. Bohr's pioneer work was the basis of a tremendous amount of subsequent work on the part of mathematical physicists, among whom the name of Sommerfeld (Germany) is outstanding. When the theory was generalized, many problems such as the complex Zeeman effect, the Stark effect, the fine structure of individual hydrogen lines, were satisfactorily explained. There were many other problems, however, such as those dealing with intensities of spectral lines and with the spectra of atoms more complicated than hydrogen, for which adequate solutions were not forthcoming. Moreover, from the outset Bohr's theory had nothing whatever to say about the actual process of emission when an electron passes from one orbit to another. So many difficulties presented themselves to theoretical physicists that revisions in theories were constantly being made, and new concepts introduced. As one example, we may mention the use of the spinning electron in order to interpret the meaning of multiple energy levels. Such work culminated in the introduction of systems of quantum mechanics in which men no longer made the picture of an atom model (such as the planetary atom of Bohr) the starting point for their theories, but rather certain abstract ideas. The most outstanding of such theories is known as wave mechanics, the foundation of which was laid by Louis

de Broglie (France). In its subsequent development the name of Schrödinger * (Germany) is outstanding.

While a detailed consideration of this work is much beyond the scope of this book, it may be stated that de Broglie, accepting the modern idea of the equivalence between mass and energy, postulated that, in every mechanical system, waves are associated with mass particles. The development of de Broglie's basic idea leads to the view that, just as optical problems relating to large openings may be solved by using the rectilinear propagation of rays of light, whereas, for small openings, the spreading out of waves is essential for a correct solution, so wave mechanics replaces ordinary mechanics when we have to deal with bodies of atomic dimensions rather than those usually encountered. Stated otherwise and perhaps more correctly, just as the wave theory leads to laws of geometrical optics for big openings but is necessary to explain diffraction effects, so the more general wave mechanics is equivalent to ordinary mechanics for bodies of large dimensions but is necessary for the correct solution of problems dealing with masses of atomic and sub-atomic dimensions. Now an essential feature of wave mechanics is that the waves associated with a particle of mass m and velocity v have a wave-length given by the relation

$$\lambda = \frac{h}{mv}. \tag{20.02}$$

Here then is a prediction of associated waves which may be tested experimentally. Let us apply it to the work of Davisson and Germer. An electron with charge of e statcoulombs which has fallen through a potential difference of V volts has acquired energy $= \dfrac{Ve}{300}$ ergs, and hence we may write,

$$\tfrac{1}{2}mv^2 = \frac{Ve}{300}, \tag{20.03}$$

where m is the mass and v the velocity of the electron. If now we eliminate v between relations (20.02) and (20.03), we obtain

* Using an entirely different method of approach Heisenberg developed a system of quantum mechanics which proved to be essentially the same as wave mechanics.

$$h\sqrt{\frac{150}{mVe}}. \qquad (20.04)$$

Substituting the standard values $e = 4.77 \times 10^{-10}$; $h = 6.55 \times 10^{-27}$; $m = 8.93 \times 10^{-28}$ gm., we have

$$\lambda = \frac{12.25}{\sqrt{V}}. \qquad (20.05)$$

Hence for 54 volt electrons, $\lambda = 1.67$ A, for 64 volt electrons, $\lambda = 1.52$ A, results in excellent agreement with the values (1.65 and 1.52) obtained experimentally on the supposition that the reflection of electron waves follows X-ray laws. It would seem, then, that good evidence has been obtained that at least sometimes electrons behave as if they were a group of waves.

182. Further confirmation of the truth of this result is provided by the work of G. P. Thomson (Great Britain), who used a method which involved the passage of a beam of high-speed electrons through thin layers of matter. Here again the experimental test made use of a standard means of measuring X-ray wave-lengths. Indeed the method was essentially the same as that used by Laue and his co-workers when they proved the wave nature of X-rays by the passage of a narrow beam of these rays directly through a crystal. Because of diffraction, with such an arrangement, a number of well-defined spots are obtained on a photographic plate receiving the emergent beam. If the crystal is replaced by a powder in which a number of small crystals are arranged at random, the spots coalesce to form a series of concentric rings, whose radii depend on the wave-length and on the crystal arrangement. If, then, electrons are accompanied by waves it is reasonable to argue that a similar result might be detained with suitable wave-lengths.

That this is actually the case, Thomson showed by the use of extremely thin layers of metal (with thickness of the order of 10^{-6} mm). By using potentials ranging from 10,000 to 60,000 volts, wave-lengths much shorter than those of Davisson and Germer

were obtained [see relation (20.05)]. When a narrow beam of such high-speed electrons was incident normally on one of these films, a photographic plate some 30 cm. away showed the presence of well-defined rings. An actual photograph of the ring system for a very thin layer of gold is given in Fig. 2, Plate VII. For this photograph the author is greatly indebted to Professor Thomson. Moreover, since the radii of the rings depend on the wave-length (and therefore on the voltage employed), as well as on the spacing of the atoms in the material, quantitative tests could be made. When calculations were made using the measured diameter of the rings, even more striking agreement was obtained between experimental results and the predictions of wave mechanics than in the work of Davisson and Germer. Again, it seems impossible not to conclude that, under certain circumstances, electrons behave like a group of waves whose wave-length is given by relation (20.02). On the other hand, it should be stated that these waves are not of the same nature as X-rays of the same wave-lengths. That this is the case will be evident at once from the following facts: (1) electron waves, as shown by Thomson, are deflected by a magnetic field; (2) the penetrating power of electron waves (which penetrate only extremely thin layers of matter) is much less than that of X-rays of equal wave-length.

Schrödinger's development of de Broglie's ideas led to the establishment of a differential wave equation which is fundamental in quantum mechanics. In this equation there appears a symbol ψ representing the amplitude of the waves associated with a particle. When this equation is applied to the case of an electron revolving about a nucleus, as in the hydrogen atom, only certain characteristic solutions are possible. These solutions correspond to discrete energy states and lead to exactly the same expression for the energy values as that given by the Bohr-Sommerfeld theory (relation 18.12). Moreover, in the course of the solution of the equation, quantum numbers n and l appear quite naturally, that is, they are not introduced in the arbitrary manner of the earlier theories. Altogether, with the application of quantum mechanics to atomic systems, it has been possible to interpret correctly the

radiation problems which could not previously be solved by the early quantum ideas.

Although the waves associated with a particle of matter are sometimes called "material," the student should not think of them as "common-sense" waves, that is, waves which can be visualized in space and in time. That they have a reality of some kind is obvious from such experimental work as that of Thomson and Davisson and Germer. But, as we have already pointed out, they must not be considered to have the same nature as electromagnetic waves. Indeed they are best described as mathematical waves governed by an equation by means of which correct predictions can be made about the behaviour of atomic systems.

183. Let us sum up, now, what we have found. (1) As far as interchanges of energy are concerned, light is a quantum or corpuscular phenomenon, whereas, its laws of propagation, as well as the facts of interference and diffraction, are accurately represented by Maxwell's equations of the electromagnetic field, which postulate a continuous wave-front. Light then seems to have a dual nature. (2) Particles such as electrons, which in a host of phenomena can be treated as mass-particles subject to ordinary dynamical laws, have been shown to act in certain circumstances, as a group of waves. Particles, then, when of atomic or subatomic dimensions, can also exhibit a dual nature. It is not amiss to think of the wave-theory as an accurate representation of facts when we have to deal with the operation of a large number of quanta, whereas in processes involving interchanges of energy where a single quantum is concerned, quantum mechanics is necessary. Similarly, Newton's laws are an exact representation of facts when we are dealing with bodies of large mass (a big aggregation of particles) but wave mechanics is necessary in dealing with a single particle of atomic dimensions.

CHAPTER XXI

CAN THE EXISTENCE OF AN ETHER BE DETECTED?

184. In Chapter XVI it was pointed out in some detail how the electromagnetic theory replaced the elastic solid theory. The resulting change in outlook did not overthrow the conception of an ether. The difficulty of imagining a medium with an elasticity sufficiently great to transmit vibrations with the speed of light, and yet so rarified as to permit apparently unrestricted motion of material bodies through it, was, of course, removed. But an ether was still considered necessary as the vehicle for the transmission of electromagnetic waves. Indeed, although the propagation of light was represented by differential equations, underlying Maxwell's work was the Faraday conception of electric tubes representing stresses in an all-pervading ether. Throughout the nineteenth century, therefore, both before and after the advent of the electromagnetic theory, it was natural that scientists should be concerned with questions relating to the relative motion of ether and terrestrial objects. Is there any experimental evidence that the earth and all objects on it are moving through an invisible medium, the carrier of light and all electromagnetic disturbances? That is the question we wish to examine in this concluding chapter.

185. Aberration of Light. — Early in the eighteenth century, an Englishman, Bradley by name, discovered a phenomenon which showed that the *direction* in which light is received from a star is altered by the motion of an observer on the earth, thus giving rise to an erroneous result in fixing the position of the star. Bradley considered his observation a proof that light has a finite velocity, and was able to calculate a value for this constant in fair agreement with the results of later and more accurate methods. Let us examine somewhat carefully this so-called *aberration of light.*

To make our ideas concrete, consider a *stationary* observer holding a long cylindrical tube with its axis AB, a, Fig. 219, in such a direction that a line of falling particles (rain-drops, for example), travelling in the direction DA pass straight down the centre of the tube. Obviously, in such a case, the direction of the axis of the tube gives the direction in which the particles are travelling.

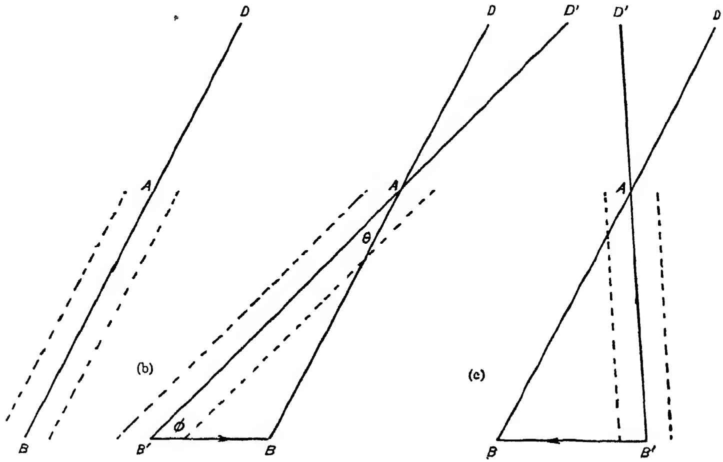


Fig. 219.

If the particles are replaced by a wave-disturbance in the direction DA, originating in some source along AD produced, the same result holds; that is, the axis of the tube gives the true direction in which the source lies.

Suppose, however, that the observer is *moving* with the tube in the direction of the arrow as in b, Fig. 219, with velocity v. In such a case, if we wish a particle entering the upper end of the tube at the point A, the centre of the opening, to emerge from the lower end at the centre, evidently the tube must be slanted so that, when the centre of the upper end is at A, that of the lower end is at B', where BB' is the distance the tube is carried along during the time that the particle travels from A to B. For, with such a

slant, the centre of the lower end of the tube reaches B at the same instant as the particle. If the falling particles are replaced by a wave-motion or by quanta coming from a source along AD produced, the direction of the axis of the tube no longer gives the direction from which the waves or quanta are coming, nor does the tube point to the true position of the source. The angle DAD' or $B'AB$, between the true direction and the axis of the tube (or the apparent direction) is called the angle of aberration.

If the cylindrical tube is replaced by a telescope receiving light from a stationary non-terrestrial source such as a star, it follows that, because of the translatory motion of the earth, the axis of the telescope does not indicate the true position of the star. Moreover, because the direction of the earth's motion continually changes throughout a year, the angle of aberration must continually change also. Thus some months after the time represented by b, Fig. 219, conditions might be represented by c, Fig. 219, where it will be noted the apparent position of the star is on the other side of the true position. Now it was just this apparent change in the position of a star which Bradley discovered, and by means of which he worked out a value for the velocity of light. Let us see how that can be done.

Let θ = angle of aberration, that is, $\angle DAD'$ or $\angle BAB'$, in b, Fig. 219, and let $\phi = \angle AB'B$, that is, the angle between the direction in which the observer is moving, and that of the axis of the telescope.

We have then,
$$\frac{\sin\theta}{BB'} = \frac{\sin\phi}{AB};$$

from which
$$\sin\theta = \frac{BB'}{AB}\cdot\sin\phi$$
$$= \frac{v}{V}\cdot\sin\phi,$$

where v is the velocity of the observer, and V that of the light.

Since θ is a small angle, we may write
$$= \frac{v}{V}\sin\phi. \qquad (21.01)$$

The maximum value of the angle of aberration is, therefore, given by

$$\theta \text{ maximum} = \frac{v}{V}.$$

Now observation gives the maximum value of θ about 20″; hence, taking v = 19 mi. per sec., we find V to be of the order of 180,000 mi. per sec.

But, it is asked, where does the ether come in, in all this work? The answer should now be clear. *Either it does not enter into the problem at all, that is, we may ignore its existence altogether; or, if it does, it must be at rest with respect to the telescope.* For if the ether were carried along with the telescope, so too would the wave-disturbance be carried along, and hence light would not continue to travel in the direction AB.

186. Airy's Experiment. — In 1871 an important variation in the aberration experiment was performed by Airy, then Astronomer Royal in England. From the explanation given above, it follows at once that the smaller V, the greater the angle of aberration. If, therefore, the air inside a telescope is replaced by water in which light travels more slowly than in air, the angle of aberration for any observed star ought to be greater. Airy tried this experiment and *found the angle exactly the same as for air.* Now, that the velocity of light in water is less than in air, there could be no doubt — experiment had shown that to be the case. There must then, it was argued, be some factor or factors which alter the velocity in the case we are considering. Moreover, since the angle of aberration remains the same, then if DA, Fig. 220, represents the direction of the light incident on the telescope, the axis of the instrument must (as before) point in the direction BAD', where $\angle DAD' = \angle BAN$ is θ, the angle of aberration. Hence if BM (equal to BB' of b, Fig. 219) represents v the velocity of translation of the telescope, AM must represent the resultant velocity of the light in the water inside the moving telescope. What, then, can bring about such a change in the velocity inside the telescope? In the first

place, it is well to note that because the incident beam of light does not strike the telescope objective normally, there will be refraction.

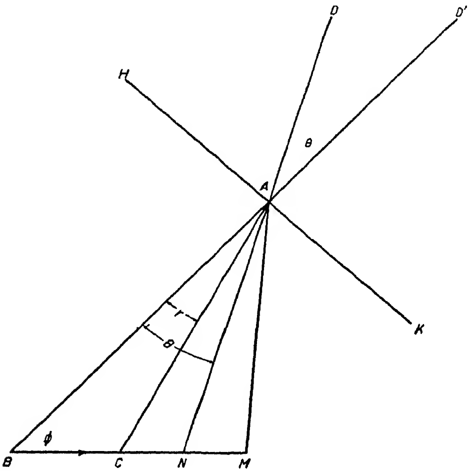


FIG. 220.

For this reason alone, the light inside the tube would travel in the direction AC with a velocity equal to V/n, the exact direction being given by the ordinary sine law,

$$\frac{\sin \angle DAD'}{\sin \angle BAC} = n \quad \text{(since } D'A \text{ is normal to } HK, \text{ the telescope objective)}$$

or

$$\frac{\sin}{\sin r} = n,$$

from which, since these angles are both small, we have

$$r = \frac{\theta}{n}, \quad \text{where} \quad r = \angle BAC.$$

The actual direction of the light in the tube, however, is along AM. An alteration in direction from AC to AM would take place if the medium carrying the light disturbance, that is, the ether, was not at rest with respect to the telescope but was carried

along by it with just the necessary velocity to allow us to represent the resultant velocity by the line AM. For this to be the case, the necessary velocity of the ether with respect to the telescope, the ether *drift* we shall call it, would necessarily be represented by CM.

Let us now find the magnitude of the ether drift, bearing in mind that BM represents a velocity v; AC a velocity V/n; CM the velocity of the ether drift; that the angle $BAC = r = \theta/n$. Let also the angle $ABM = \phi$.

For $\triangle\ BAC$ (Fig. 220), we can write

$$\frac{\sin r}{\sin\phi} = \frac{BC}{AC} = \frac{nBC}{V}$$

$$\therefore \frac{\theta}{n\sin\phi} = \frac{nBC}{V}$$

$$\frac{\quad}{\sin\phi} \qquad \frac{n^2BC}{V}. \tag{21.02}$$

For $\triangle\ BAB'$, b, Fig. 219,

$$\frac{\sin}{\sin\phi} \quad \frac{BB'}{AB} = \frac{v}{V}$$

$$\sin\phi \qquad \frac{v}{V}. \tag{21.03}$$

Hence, combining relations (21.02) and (21.03), we have

$$\frac{n^2BC}{V} = \frac{v}{V}$$

or $$BC = \frac{v}{n^2}.$$

But $$BC = BM - CM$$

$$= v - \text{ether drift},$$

$$\text{velocity of ether drift} = v - BC$$

$$= v - \frac{v}{n^2}$$

$$= \left(1 - \frac{1}{n^2}\right)v. \tag{21.04}$$

Note what has been proved. The observed result that there is no change in the angle of aberration when air is replaced by water, would follow if the ether were dragged along with the water of the telescope by $\left(1 - \frac{1}{n^2}\right)$ of the velocity of the instrument. Taking $n = 1.33$, we find this fraction has the value 0.44. For air, where $n = 1$ (approximately) this result implies no ether drift, in line with the conclusion already noted. Now the remarkable feature about all this is that an ether drift of exactly this magnitude had been deduced years before by Fresnel on the supposition that the density of the ether varies with the material body in which it is found. Moreover, over a decade before the work of Airy, Fizeau (France, 1819–1896) had performed another celebrated experiment in which he showed that the interference fringes obtained by superimposing two beams travelling equal paths through columns of moving water, one beam with the stream, the other against, were shifted from the position of such fringes when the water was stationary. Granted an ether drift, this result at once follows. From the fringe shift observed with a given velocity of water, it is not difficult to calculate the magnitude of the ether drift. The results obtained not only by Fizeau, but over thirty years later, with more refined apparatus, by Michelson and Morley verified Fresnel's formula. It would seem, then, *as if* evidence had been obtained that ether is dragged along by moving matter. To this point reference will again be made. Before leaving it, it is well to note the following:

187. (1) In both the aberration experiment and in that of Fizeau, there is *relative* motion between the instrument and some other material object. Thus the telescope moves with respect to the star, and, in Fizeau's experiment, the water moves with respect to the instrument receiving the light.

(2) If we accept as established the existence of the ether drift of an amount given by Fresnel's coefficient, it follows that, as far as first order effects are concerned, it is impossible to detect any relative motion between the *earth* and the ether. The reason

should be clear from a consideration of the following experiment. Suppose a set of mirrors A, B, C, D, is arranged as in Fig. 221, so that interference fringes are obtained between rays which have been reflected from mirror to mirror in a clockwise direction and those reflected in a contra-clockwise direction. Let L represent the length of a glass block in the path of one arm, of total length l, of such an interferometer, and suppose the apparatus so placed

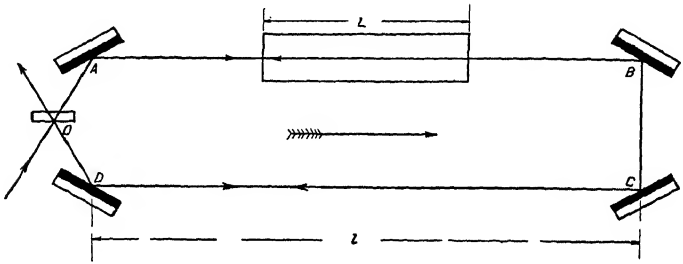


FIG. 221.

that the earth has a velocity v in the direction of the arrow. Inside the glass block the ether then has a velocity $\left(1 - \frac{1}{n^2}\right) \cdot v$ or ζv, we shall write it for short, where $\zeta = 1 - \frac{1}{n^2}$.

In finding the difference in the times taken by the clockwise and the contra-clockwise beams to cover the complete path $OABCDO$, we shall make use of the fact that the motion of an object through a stationary ether is exactly the same as if the object were stationary and the ether streaming with equal speed in the opposite direction — a virtual ether stream we shall call it. In other words we may superimpose on the whole system a velocity equal and opposite to v. Hence if T_1 = the time taken by the clockwise beam, we may write

$$T_1 = \frac{l - L}{V - v} + \frac{L}{\frac{V}{n} - (1 - \zeta)v} + \frac{l}{V + v}$$

$$\frac{l-L}{V-v}+\frac{L}{\frac{V}{n}-\frac{v}{n^2}}+\frac{l}{V+v}, \tag{21.05}$$

since V = velocity in air for a stationary ether,

$V - v$ = velocity against the virtual ether stream,

$\frac{V}{n}$ = velocity in glass for a stationary ether, and

$(1 - \zeta)v$ = the virtual ether stream in the glass (v and ζv in opposite directions).

Similarly T_2, the time taken by the contra-clockwise beam is given by

$$T_2=\frac{l-L}{V+v}+\frac{L}{\frac{V}{n}+\frac{v}{n^2}}+\frac{l}{V-v}. \tag{21.06}$$

Now if these two times are subtracted, it can easily be shown by elementary algebra that when the ratio v^2/V^2 is neglected in comparison with unity, the difference is zero. In other words, if it were sought to detect a shift in the interference fringes observed with such an arrangement when the apparatus was rotated through 180°, a negative result should be obtained. *Such is actually the case.*

If, however, the factor v^2/V^2 is retained, T_2 is not exactly equal to T_1. Hence, if fringes were observed for the position indicated by Fig. 221, and the apparatus were then rotated through 180°, at any given place in the field of view where the two beams meet, there would be a phase difference introduced corresponding to a time difference of $2(T_2 - T_1)$. Theoretically, therefore, there ought to be an observed shift; but whether it could be observed or not, would depend on its magnitude. With the apparatus we have been discussing, as already noted, actually no shift can be detected. Is it possible to do so with any arrangement? If so, the apparatus must be sufficiently sensitive to detect a difference of about 1 part in 100,000,000.* Because of the shortness of light waves such

* Taking v = 19 mi. per sec.; V = 186,000 mi. per sec.; $\frac{v}{V} = 10^{-4}$ approximately. Hence $\frac{v^2}{V^2} = 10^{-8}$.

an apparatus can be designed, and hence a test made to see if relative motion between the earth and the ether can be detected even when second order quantities are taken into consideration.

188. The Michelson and Morley Experiment. — Probably the most famous of such tests was the pioneer one made by Michelson and Morley in 1881. The arrangement utilized was essentially a Michelson's interferometer, with each optical path lengthened by

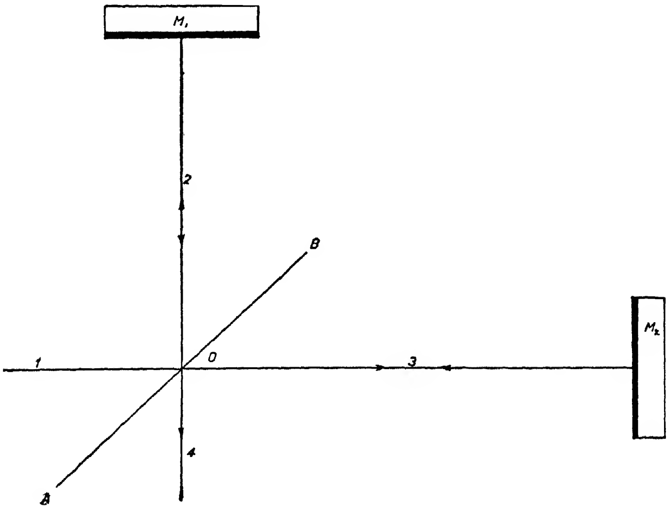


FIG. 222.

the use of several mirrors reflecting the light back and forward, while the test consisted in watching for the second order shift in interference fringes which should be observable *if* relative motion of the earth and the ether could be detected. To understand the nature of the test, consider a simple interferometer arrangement such as represented in Fig. 222, where ray 1 is partly reflected, partly transmitted at the half-silvered surface AB, thus giving rise to rays 2 and 3, which after reflection at mirrors M_1 and M_2

subsequently come together in direction 4. Suppose that the optical paths of rays 2 and 3 are each equal to l, and that the interferometer is so placed that ray 3 is parallel to the orbital motion of the earth. As has already been pointed out, if the instrument is moving through a stationary ether with velocity v, the problem is the same as if the instrument were stationary, while an ether streamed past with velocity v in the opposite direction.

If, then, V = the velocity of light in a stagnant ether,
$V + v$ = the velocity with such an ether stream,
$V - v$ = the velocity against such a stream.

Hence, the time required for light to go from O to M_2 and back to O

$$\frac{l}{V - v} + \frac{l}{V + v} = \frac{2lV}{V^2 - v^2}$$

$$\frac{2l}{V\left(1 - \frac{v^2}{V^2}\right)} = \frac{2l}{V} \cdot \left(1 + \frac{v^2}{V^2}\right), \qquad (21.07)$$

if we neglect the fourth and higher powers of $\frac{v^2}{V^2}$.

To find the time required to travel from O to the mirror M_1 and back again, it is well to realize that the problem is essentially the same as that of a boat crossing a flowing stream. In such a case it is necessary to know the resultant velocity in the direction of the path. In the direction OM_1 we see at once from the vector diagram of Fig. 223, that the velocity has the magnitude $\sqrt{V^2 - v^2}$. Hence the time to travel the path OM_1

$$= \frac{l}{\sqrt{V^2 - v^2}}.$$

Since an equal time is taken to cover the path M_1O, the total time

$$2l \qquad \frac{2l}{V\left(1 - \frac{v^2}{V^2}\right)^{\frac{1}{2}}}$$

$$= \frac{2l}{V}\left(1 + \frac{v^2}{2V^2}\right), \qquad (21.08)$$

again neglecting the fourth and higher powers of $\frac{v^2}{V^2}$.

Hence ΔT, the *difference* in the times for the two paths is given by

$$\Delta T = \frac{2l}{V}\cdot\left(1 + \frac{v^2}{V^2}\right) - \frac{2l}{V}\cdot\left(1 + \frac{v^2}{2V^2}\right)$$

$$= \frac{lv^2}{V^3}. \tag{21.09}$$

Suppose, now, that the whole apparatus is carefully rotated through 90° so that the path OM_1 becomes parallel to the direction

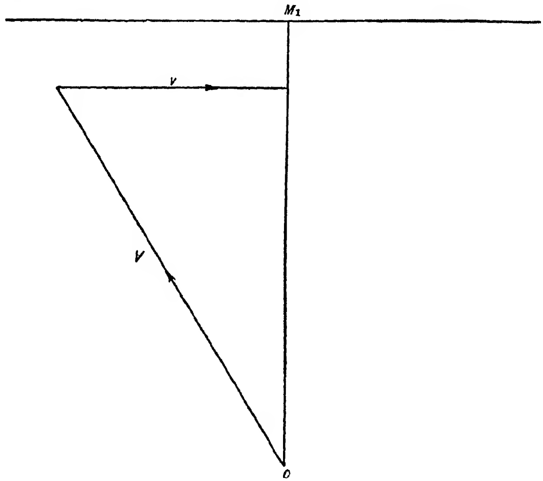


FIG. 223.

of the earth's motion. Such a rotation then introduces a difference in time $= \frac{2lv^2}{V^3}$. That is, at any given place in the field of view where fringes are observed, a difference of phase is introduced of an amount corresponding to this time difference, hence a fringe shift should be observed provided the magnitude is great enough. Let us see what shift is to be expected in the Michelson and Morley arrangement.

If f = the fringe shift,
λ = the wave-length of light used,

then the time difference $= f \cdot \dfrac{\lambda}{V}$.

$$\therefore\ f \cdot \frac{\lambda}{V} = \frac{2lv^2}{V^3}$$

or

$$f = \frac{2l}{\lambda} \cdot \frac{v^2}{V^2}.$$

In one actual arrangement, l was about 1100 cm., while $\dfrac{v}{V}$, as has been already pointed out, $= 10^{-4}$ approximately. Therefore,

$$\text{for } \lambda = 6 \times 10^{-5} \text{ cm., we have}$$

$$f = \frac{2 \times 1100}{6 \times 10^{-5}} \times 10^{-8} = 0.4.$$

If, therefore, such an instrument is moving through a stationary ether, on rotation through 90° a shift of about 0.4 of a fringe, *one readily discernible*, should be observed. "In order to minimize the displacements due to external factors . . . the interferometer was mounted on a block of stone 1.5 m. square and .25 m. thick resting on an annular wooden ring which floated the whole apparatus on mercury" (Michelson). But in spite of the most elaborate precautions, *no shift was observed.*

In 1905 the Michelson and Morley experiment was repeated by Morley and Miller, with more sensitive apparatus, and the same negative result was obtained. Moreover, other methods of detecting a relative motion of the earth and the ether were tried, and these, too, led to the same conclusion. Into the details of other methods we shall not enter, except to indicate that in one well-known experiment use was made of the fact that, if this relative motion existed, a couple should act on a parallel plate electric condenser delicately suspended. In 1902 Trouton and Noble (Great Britain) made a test of this sort with a negative result, a conclusion supported by the recent repetition of their work with greater refinement by Tomaschek (Germany), in 1925, and by Chase (United States), in 1926.

During the years 1921 to 1925, an elaborate series of observations was carried out by Miller, the result of which seems to indicate an ether drift with a velocity about one-third of the orbital motion of the earth. This drift, however, is not the result of the earth's orbital motion but rather is considered (if a real effect) to be due to "a velocity of the solar system relative to stellar space" (Michelson) or, to quote Miller himself, to be "a systematic cosmical effect as of a true ether drift." But even if Miller's recent experiments (together with all others of its kind) indicate no ether drift as a result of the earth's orbital motion, his work has consequences which cannot be ignored, for it affects the whole basis of the theory of relativity. (See below.) For that reason, other repetitions of the ether drift experiments have been made within the last two or three years, the utmost care being taken to make such tests as refined as possible.

Of these, reference has already been made to that due to Tomaschek and that to Chase. Another important test was made by Kennedy at the California Institute of Technology with an apparatus capable of detecting a shift one-quarter of that observed by Miller. The result was negative. Michelson himself, assisted by Pearse and Pearson, have repeated the original Michelson and Morley experiment with somewhat different conditions, and with the most elaborate precautions, and once again the original negative result was obtained. *The weight of the evidence is then decidedly in favor of the view that the existence of an ether cannot be proved by such experiments*, and yet Miller's results cannot be ignored. It may be that there are flaws in respect to the theory underlying the use of the various instruments. In any case it is well for the student not to take too dogmatic a position on this (or on any other) question but to recognize that the problem has not yet been entirely settled.

188A. In 1935 there is little change in the position of this question. In the *Review of Modern Physics* for July, 1933 Miller has a lengthy discussion of the whole question, with no essential change in his conclusion. On the other hand Joos (Germany) has

repeated the Michelson-Morley experiment and obtained a negative result. And modern theoretical physics, so much of which has for its foundation this negative result, has added to its triumphs in its ability to explain and to predict facts.

189. For years before Miller's disquieting results were announced, many physicists had concluded that *it is not possible to detect a motion of the earth with respect to an all-pervading ether.* Accepting that conclusion, scientists had then to face a very important question. If there is an ether through which the earth is moving, why then the failure to detect a fringe shift which could readily be observed? This was the question the answer to which was much debated after the original Michelson and Morley experiment. To it three answers may be given.

(1) The premise is wrong, an ether does not exist. It will be pointed out presently that the theory of relativity explains *all* the phenomena discussed in this chapter without postulating an ether, so that this is indeed a possible solution. To this point we shall return.

(2) The ether is of such a nature that in the neighborhood of the earth's surface it is dragged along with the earth, there thus being no relative motion of the two. For a time this idea was given serious consideration although it led to difficulties concerning the nature of the ether. It was abandoned, however, because an apparently satisfactory explanation was provided by the third answer to our question — one, moreover, which involved ideas ultimately shown to agree with other facts not related to the particular problem we are discussing.

(3) **The Fitzgerald-Lorentz Contraction.** — An *ad hoc* explanation (that is, one especially devised to fit the case) which proved to have a wider scope than its author realized was suggested by Fitzgerald (Great Britain). According to his hypothesis, the earth moves through a stagnant ether, but all bodies on its surface suffer a contraction in the direction of their motion relative to the ether, none at all in a direction at right angles. If this change in length is in the ratio $1 : \sqrt{1 - \frac{v^2}{V^2}}$, it is not difficult to show

that the dimensions of the Michelson and Morley apparatus would so alter that no phase difference is introduced on rotation through 90°, hence no fringe shift. By actual measurement, any such change in dimension could never be demonstrated, for the dimensions of rulers and all measuring instruments would always alter in the same ratio as those of the bodies to which they were applied. The Fitzgerald hypothesis, therefore, being of the *ad hoc* variety, and, indeed, being one which raised certain serious difficulties which need not here be considered, could not by itself be considered a highly satisfactory solution of the problem.

By means of an extension due to Lorentz (Holland), however, the same contraction formula was deduced on a more fundamental basis. Moreover, the hypothesis of Lorentz led to results confirmed by experiment. In brief, by utilizing standard electromagnetic theory, Lorentz showed that, *granted a contraction of individual moving electrons*, the bulk contraction of Fitzgerald followed. Stated thus, there may seem little difference in the two theories, but, when it is realized, not only that Lorentz' work did not give rise to the difficulties encountered in the original proposal, but also that it led to a conclusion verified by direct experiment, it will be realized how much more fundamental was this new viewpoint. According to the Lorentz theory, the *mass* of an electron

should increase with its speed in the ratio 1:

quent experiments on high-speed electrons amply verified this conclusion. It would seem, then, as if a satisfactory explanation of the negative result of the Michelson and Morley experiment had been provided, — one, moreover, which accepted the idea of an ether with reference to which the absolute motion of bodies could be expressed.

190. But the negative result was to have consequences that were much more far-reaching than those of the Fitzgerald-Lorentz contraction hypothesis, for it led to the theory of relativity, one of the most remarkable developments of modern physics. While the author feels it is wise to conclude this book by leaving the

reader at the entrance to this fascinating field, and therefore has no intention of going into the details of relativity, one or two general remarks are not amiss. As most people know, the foundation to this theory was laid by Einstein (Germany), who, as a basis on which to build his structure, postulated two things.

(1) The velocity of light is independent of the velocity of the source. About this there was nothing new, for it is confirmed by observation, and, indeed, for that reason, it is a fact rather than a postulate.

(2) It is impossible to detect a uniform motion of the earth or any system through space by observations made on that system. In effect this says, there is no such thing as absolute motion relative to an all-pervading ether. No experiments have ever shown the existence of such a motion and none ever will. Accepting this postulate, further experiments such as the Michelson and Miller are, of course, useless.

With these postulates as a basis, the (special) theory of relativity shows that the dimensions of moving bodies vary with their velocities in accordance with the law exemplified by that governing the Fitzgerald-Lorentz contraction; that the mass of a moving body depends on its velocity subject to the law given by Lorentz; and, in fact, that all the phenomena discussed in this chapter find a ready interpretation. For all the work on relativity, the ether concept is superfluous. Is, then, the ether to be discarded? It may be so. On the other hand, it is difficult for many minds to throw overboard the idea of an ether as the vehicle necessary for the transmission of electromagnetic energy, just as it is difficult for many to be content with a differential equation as a complete description of a wave-motion. But many of our fundamental concepts are in the refining pot, and it is rash to predict what the future has in store. Regarding the ether and the theory of relativity, Michelson is not alone in holding the opinion that "it is to be hoped that the theory may be reconciled with the existence of a medium, either by modifying the theory or, more probably, by attributing the requisite properties to the ether." Even Eddington

(Great Britain), the well-known relativist, does not discard the ether. According to him, "we postulate ether to bear the characters of the interspace, as we postulate matter or electricity to bear the characters of the particles."

191. In this concluding paragraph we shall make a brief reference to a law deduced by the theory of relativity which has amazing possibilities. According to this law there is an equivalence between mass and energy, expressed by the relation $E = mc^2$, where E is the quantity of energy resident in a mass m, c is the velocity of light in free space. Now, in modern work dealing with the structure of matter, there is much evidence (1) that the nuclei of all atoms are made of electrons and protons (a proton being the name given the hydrogen nucleus); (2) that in the coming together of protons and electrons to form a nucleus there is a *loss* in mass. Consider one example, that of helium. From its atomic weight, 4.00216, it is evident that 4 hydrogen nuclei or protons must come together (along with 2 electrons) to form the helium nucleus. But the mass of 4 hydrogen nuclei $= 4 \times 1.00778$ or 4.03112. There is, therefore, a loss of mass $= 4.03112 - 4.00216$ or 0.029 when a single helium atom is thus formed.

According to this relativity law, an equivalent amount of energy must replace this "lost" mass. Is the energy which is thus created, radiation? If so, the student cannot fail to realize the tremendous possibilities of such a transformation. For example, if, as the result of the formation of one helium atom, all the radiant energy thus formed went into a single quantum, it is not difficult to show that the frequency of the radiation would be extremely high, or the wave-length would be much shorter than the shortest gamma rays. Have we here, then, the origin of one part of the so-called cosmic rays? Again, astrophysicists are suggesting that one of the sources of the tremendous stellar output of radiation is found in the loss of mass when a proton and an electron come together and are annihilated. The energy-mass law suggests, too, the possibility of the converse process, the creation of electrons and of protons out of radiant energy. If the contemplation of

such questions makes the reader realize that this book has been, indeed, but an introduction to the study of light and if the reader is left with a desire to go more deeply into the subject, the author will feel that the time devoted to its writing has not been misspent.

TABLE XIV

Some Useful Wave-Lengths in the Laboratory

Source	Wave-Length
Hydrogen vacuum tube	6562.79 (H_α) 4861.33 (H_β) 4340.47 (H_γ) 4101.74 (H_δ)
Helium vacuum tube	7065.19 6678.15 5875.62 — (D_3) 5015.67 4921.93 4713.14 4471.14 4387.93 4026.19 3888.65
Mercury vapour lamp (glass)	5790 5770 5460.7 4358.3 4046.5
(quartz)	$\left.\begin{array}{l}3662.9\\3654.8\\3650.1\end{array}\right\}$ $\left.\begin{array}{l}3131.5\\3125.7\end{array}\right\}$ 2967.3 2536.5
Sodium flame	5895.93 — (D_1). 5889.96 — (D_2).

TABLE XV

A few Members of the Principal, Sharp and Diffuse Series of Lithium.

(Wave-lengths from Fowler's Report on Series in Line Spectra.)

Principal 1 S − mP.			Sharp 1 $P - mS$		
1 S = 43486.3			1 P = 28582.5		
	m	mP		m	mS
6708.85	1	28582.5			
3232.61	2	12560.4	8126.52	2	16280.5
2741.31	3	7018.2	4971.93	3	8475.2
2562.50	4	4473.6	4273.28	4	5187.8
2475.29	5	3099.2	3985.79	5	3500.4

Diffuse 1 $P - mD$		
1 P = 28582.5		
	m	mD
6103.53	2	12203.1
4602.99	3	6863.5
4132.29	4	4389.6
3915.0	5	3047.0

ANSWERS TO PROBLEMS

Chapter II (p. 54)

1. 83.3 cm. **2.** above, below, at, below, normal level. **4.** (a) 0, $+ 6.9$; (b) 0, $- 6.9$. **6.** 12.1 units. **7.** 10.4. **8.** $45°$. **11.** $x = a \sin \frac{2\pi t}{T}$, $y = a \cos \frac{2\pi t}{T}$. **12.** $120°$. **13.** -8. **14.** $+ 6$. **15.** $- 4.2$ cm. **17.** (a) $60°$, (b) $- 3.86$ cm. **21.** (a) 40 cm. per sec., (b) $60°$, (c) $315°$, (d) 0.79 or $- 3.8$. **22.** (a) 0.88, (b) $60°$, (c) $- 3.65$. **23.** 5 cm. **24.** (a) 12, 12.8, $- 6$, (b) no, (c) 4,800 cm. **25.** (a) 10.1, (b) $y = 10.1 \sin \left(\frac{2\pi t}{T} - 1° \, 30'\right)$.

Chapter III (p. 94)

1. $9''$. **2.** 225×10^6 m. per sec. **4.** 6,666.7. **5.** 225×10^6 m. per sec. **6.** $52° \, 48'$. **7.** 3.73 cm. **8.** (b) $\sin^{-1}\frac{V}{v}$. **10.** 0.74 cm. **12.** $+ 80$ cm. **13.** 1.5. **15.** $- 133$ cm. **16.** $+ 17.1$ cm. **17.** $+ 52$ cm. from lens. **18.** $+ 75$ cm. from face. **19.** $- 5$ cm. from lens. **20.** $- \frac{r}{n - 1}$. **21.** $- \frac{\;}{n - 1}$ **23.** 20 cm. from curved surface. **24.** $+ 35$ cm.

Chapter IV (p. 115)

1. 49. **2.** 3.8 cm. **3.** (a) 9.16 cm., (b) 13.3 cm. **4.** $+ 30$ in. from converging lens. **5.** (a) in same position as flame, (b) 10 times object, (c) virtual. **6.** (a) $+ 10$ cm. from second lens, (b) 3.3 cm. **7.** $- 240$ cm. from second lens, (b) 30 cm. **8.** 6 cm. in front of mirror. **9.** (a) $- 10$ cm. from converging lens, (b) 12 cm. **10.** (a) $- 24$ cm. from diverging lens, (b) 0.8 cm. **11.** (a) $+ 9.4$ cm. from second lens, (b) 1.25 cm. **12.** $- 60$ cm. from first lens. **13.** $+ 30.0$ cm. from second surface. **15.** (a) $\alpha = + 0.336$ cm., $\beta = - 0.336$ cm., (b) $+ 33.3$ cm. from second surface, (d) $+ 33.1$ cm. from centre of lens. **16.** $\alpha = + 0.54$ cm., $\beta = - 0.81$ cm. **17.** $+ 23$ cm. from second surface.

Chapter V (p. 145)

1. 90 cm. and 10 cm. **2.** 6.25 to 7.25; 6.25. **3.** (a) + 5.8 cm. from eyepiece, (b) 5 mm., (c) $3° 16'$. **4.** 1.5 cm. **5.** 95.7 m., 5.1 cm. **6.** 6 cm. **7.** (a) − 1.477 cm. from first lens, (b) 2.1 cm. **8.** 7.4.

Chapter VI (p. 169)

1. 4.57×10^{14}. **2.** $39° 55'$ with face. **3.** 1.596. **4.** 1.621. **5.** $65° 33'$, $67° 24'$. **6.** 1.597. **7.** (a) $48° 24'$, $49° 40'$, (b) 0.68 cm. **8.** 1.5 mm. **9.** 6.5 cm. **10.** (a) $44° 40'$, (b) 1.55 cm. **14.** $51° 14'$. **15.** 1.33 mm. **16.** $7° 46.7'$; $1.21°$. **17.** (a) $5° 8.6'$, (b) 1.252, 1.250, 1.253, 1.249, 1.250; (c) $5° 6.7'$. **18.** (a) $2.2'$, (b) 0.4 mm. **19.** $12.8°$ (crown), $8.2°$ (flint). **20.** diverging. **21.** (a) 3 mm., (b) 64.0 cm. **22.** − 28.3 cm.; 68.4 cm. **23.** + 11.3 cm., + 23.6 cm. **24.** (a) 10.24 cm., − 14.5 cm., (b) + 10.30, + 10.17, + 9.92; −14. 59, − 14.25, − 13.85. **25.** (a) 81.6 cm., (b) +23.6 cm., − 33.21 cm. **26.** 1.557. **27.** + 43.7 cm., − 24.3 cm. **28.** 20 cm.

Chapter VII (p. 204)

1. 6.2 km. per sec. **2.** 506.6 (taking 1100 ft. per sec. as velocity of sound in air). **5.** (a) λ 9444 A, (b) 15,980, (c) 17.41×10^3. **6.** see Table VIII.

Chapter IX (p. 247)

1. 5.9×10^{-5} cm. **2.** 6679.99 A. **3.** 1.33 mm. **4.** 6×10^{-4} cm. **5.** fringes 6 mm. in width. **6.** 1.5. **7.** 0.915. **8.** 0.75. **10.** (b) 6×10^{-5} cm. **12.** 0.9 mm. **13.** 2.25 mm. **14.** 0.000736 cm., 679 cm. **15.** 0.866 cm. **16.** 2.54 mm. **17.** 1.44. **18.** 0.0109 mm. **19.** 6×10^{-5} cm. **20.** 0.00589 cm. **21.** 0.289 mm.

Chapter XI (p. 308)

2. 0.707 cm. **3.** 50 cm. from zone plate. **4.** 50 cm. (or 16.7 cm. or 10 cm.). **5.** 6×10^{-5} cm. **6.** (b) 0.14 mm. (one-third first half-period element gives a resultant amplitude equal to $\frac{d_1}{2}$.

7. In (1) centre is always bright. **8.** As in Fig. 6, Plate V. Dark bands on either side of centre 0.6 mm. apart. **9.** 15 mm. **10.** As in Fig. 6, Plate V. Centre to first dark band 0.9 mm. **11.** 0.13 secs. **12.** (1) 0.15 mm., (2) 13 min., (3) 1.1 mm. **13.** (c) 3.8 min., (d) as in Fig. 6, Plate V. **15.** 0.0125 mm. **16.** 0.88 mm. **17.** (1) As in Fig. 6, Plate V, (2) about 6.9 min. **19.** 9.6 mm. **20.** Central image with several both horizontal and vertical orders 2.5 mm. apart. **21.** 5.88×10^{-5} cm. **22.** 4:5. **24.** 6×10^{-5} cm. **25.** (a) 17° 16′, (b) 3 cm. **26.** 5.89×10^{-4} mm.

Chapter XV (p. 374)

4. 3:2. **5.** 4 images with intensities in ratio 1:3:1:3. **6.** 2:3. **7.** (b) 1:0.75. **10.** 0.011 mm. **11.** (1) phase difference an even multiple of π; (2) odd multiple of π. **12.** 0.016 mm. **14.** (a) 0.016 mm. **15.** (a) 1:3; (b) 0.033 mm. **16.** 1:3. **17.** 0.9π. **21.** 4.5 mm.

SUPPLEMENTARY PROBLEMS

1. At a certain instant the *shape* of a simple train of plane waves is given by

$$y = 6 \sin \frac{\pi x}{100}.$$

If the velocity of the wave disturbance is 40 cm. per sec., in a direction away from the origin, find the equation giving the shape 1 second later.

2. Plane waves travel along the x-axis in the positive direction at 10 cm. per second and cause the origin to move subject to the law

$$y = 6 \sin \frac{2\pi t}{10}.$$

Derive the equation giving the *shape* of the waves at the instant $t = 10$ seconds.

3. When plane waves of length 240 cm. traverse a medium, at a certain instant a particle A has a displacement of $+4.0$ cm. and the displacement is increasing. Find (i) the amplitude of the disturbance if at this instant the phase angle of A is $45°$; (ii) the displacement and phase angle of a particle B 40 cm. beyond A, at the same moment.

4. Plane waves are traversing a line of particles along an x-axis in the positive direction, with amplitude 6 units. At a certain instant the displacement of a particle A 20 cm. from the origin is $+4$ units (and increasing). At the same instant the displacement of a particle B 30 cm. from the origin is $+3$ units (and increasing). Find the wave-length of the wave-motion. (Assume source at origin vibrates subject to $y = 6 \sin \frac{2\pi t}{T}$.)

5. A particle is simultaneously acted on by the following three S.H.M.

$$y_1 = 3 \sin \frac{2\pi}{T}\left(t - \frac{T}{12}\right); \quad y_2 = 4 \sin \frac{2\pi}{T}\left(t + \frac{T}{8}\right);$$

$$y_3 = 5 \sin \frac{2\pi}{T}\left(t - \frac{T}{6}\right).$$

Find graphically, by rule and protractor, *without* plotting graphs, (i) the resultant amplitude, (ii) the equation of the resultant motion.

6. A S.H.M. plane wave disturbance of amplitude 6 units and periodic time 0.5 sec. travels along an x-axis in the positive direction at the rate of 200 cm. per second.

(a) If, at a certain instant, the phase angle of a particle A is $+15°$, what is the phase angle, at the same instant, of a particle B, (i) 30 cm. farther away from the source than A, (ii) 20 cm. nearer the source?

(b) What is the wave-length of the disturbance?

(c) What is the equation giving the shape of the waves at a moment such that the displacement of a particle at $x = 0$ is zero (and approaching the side of positive displacement)?

7. Plane waves of amplitude 12 units and period 10 sec. travel in a certain direction at the rate of 40 cm. per sec. At a certain instant a particle A has a displacement of $+6$ decreasing. What is the displacement of a particle B 100 cm. from A (in the direction along which disturbance is travelling)?

8. Plane waves of amplitude 6 units, arising from a S.H.M. disturbance, travel along an x-axis in the positive direction. At a certain instant a particle A has a displacement of $+3$ units (decreasing) and at the same instant a second particle B 31.5 cm. *farther away* from the source has a displacement of -1.571 (increasing in a negative direction). Find (i) the phase angle of A; (ii) the phase angle of B; (iii) the wave-length of the motion (more than one value is possible); (iv) the displacement of a particle 9 cm. *nearer* the origin than A.

9. At a certain instant the shape of a train of plane waves travelling along an x-axis in the positive direction is given by

$$= 6 \sin \frac{2\pi x}{400}$$

After $\frac{1}{2}$ second, the shape is given by $y = 6 \sin \left(\frac{2\pi x}{400} - \right.$
Find the velocity of the wave disturbance.

10. A particle is subject simultaneously to the following two S.H.M.

$$y_1 = 4 \sin \frac{2\pi t}{8}; \; y_2 = 6 \sin \left(\frac{2\pi t}{12} - \frac{\pi}{3}\right).$$

Find, by using a graphical method, (i) after what time the resultant displacement is first equal to zero; (ii) the period of the resultant; (iii) the resultant amplitude.

11. An object is viewed through a combination of a converging lens of focal length 30 cm. and a diverging lens of focal length 5.70 cm. placed 30 cm. apart. The eye is placed at a point on the axis just behind the centre of the diverging lens, and the object is 180 cm. from the converging lens. Find where the final image is.

12. A small bright source of light is placed at a point on the axis 30 cm. from the centre of a thin converging lens of focal length 15 cm. 20 cm. beyond the converging lens is a coaxial diverging lens of focal length 20 cm. If the aperture of the converging lens is 4 cm. and that of the diverging at least as great, find the diameter of the patch of light received on a screen placed at right angles to the axis of the lens system at a distance of 10 cm. from (and beyond) the diverging lens.

13. A small source is placed at a point on the axis 20 cm. from the centre of a thin diverging lens of focal length 20 cm. Where must a concave reflecting mirror of radius of curvature 30 cm. be placed (beyond the lens) so that the rays of light retrace their path to the source after reflection?

14. When a converging glass lens is placed near the centre of a long trough (which is at first empty) a real image of a small electric bulb 100 cm. from the lens is formed 50 cm. from the lens. If the tank trough is now filled with water, where will the image be?

Index for air-glass = 1.50; for air-water = 1.33.

15. A small object is placed 100 cm. from a *concave surface* of radius of curvature 20 cm. separating air from glass (index 1.6). Find, by method of change of curvature, where the image is.

16. In a certain simple telescope the objective has a diameter of 20 cm. and a focal length of 200 cm. What must the focal length of the eye lens be so that a pupil of diameter 3 mm. is just completely filled with light when placed at the eye-ring? Illustrate your solution with a careful diagram.

17. A Ramsden eye-piece is used as a magnifying glass with an object so placed that parallel rays emerge. If the common focal length of each lens is 6 cm. and the aperture of the lens next the object is 1 cm., find the largest size of object which can be seen with the eye *immediately* behind the centre of the eye-lens.

18. Two thin lens, A, diverging of focal length 20 cm., and B, converging of focal length 15 cm., are placed coaxially with their centres 5 cm. apart. A very bright small object is placed at a point on the axis 15 cm. from A (and 20 cm. from B). If the aperture of each lens is 4 cm. find the diameter of the patch of light received on a screen placed 20 cm. from B (and 25 cm. from A) with its plane perpendicular to the axis.

19. An object is placed 60 cm. from a converging lens of focal length 30 cm. and aperture 4 cm. Find the size of the largest object all parts of which can be seen (not necessarily with equal brightness) by an eye (treated as a point) placed on the axis at a point 90 cm. beyond the lens.

20. A glass double-convex lens whose faces have radii of curvatures of 20 cm. and 30 cm., is bounded on one side by air; on the other by water. If parallel rays are incident on the air side, find, by the method of change of curvature, where an image is formed.

Index: air-glass = 1.60; air-water = 1.33.

21. A glass double-convex lens of negligible thickness, each of whose faces has a radius of curvature of 20 cm., floats on clean mercury. Find at what distance from the lens a small object must be placed so that the resultant image (arising from reflection at the mercury) coincides with the object.

Index: air-glass = 1.50.

22. A small object is viewed through a combination of two converging lenses placed 2 cm. apart, the first (next the object) of focal length 4 cm., the second (next the eye) of focal length 6 cm. If the object is 2 cm. from the first lens, find the *magnifying power* of the combination.

23. When a small object is viewed through a magnifying glass of focal length 6 cm. the image is formed 30 cm. from the lens. Find the *magnifying power*.

24. A small object is viewed through a simple magnifying glass. If the object is 5 cm. from the glass find the *magnifying power*. What limitation is placed on the focal length of the glass?

25. A spectrograph has two equilateral 60° prisms, each of which is set in minimum deviation for a wave-length whose index for each prism is 1.632. Find the angle the bases of the two prisms make with each other.

26. A lens has a focal length of 40.2 cm. for C light; 39.8 cm. for D light; 39.0 cm. for F light. Find the value of the dispersive power of the material of which it is composed.

27. Parallel rays of monochromatic light fall on a 60° prism set in the position of least deviation and the emergent light is observed in a telescope, the cross-wires being set on the image. If the prism is rotated through 2°, through what angle will the telescope have to be rotated to keep the image on the cross-wires?

Index of prism for this light = 1.600.

28. A 60° equilateral prism has sides 2 cm. long and is placed in the position of least deviation for light whose index is 1.600. If parallel rays are incident on the whole face of the prism find the least aperture a lens can have in order to receive all the light which leaves the prism.

29. An achromatic lens combination made of a converging crown and a diverging flint has a focal length of 50 cm. If the focal length of the crown lens for C light is 20 cm., find the focal length of the flint lens for F light. Given index, for flint, for $C = 1.531$; for flint, for $F = 1.583$.

30. A narrow rectangular aperture of width 0.075 mm. on which fall plane waves of wave-length 6×10^{-5} cm. is placed in front of a converging lens of focal length 100 cm. and the diffraction pattern is observed in the focal plane of the lens. Find the ratio of the intensity of the light at the centre of the pattern to that at a point 4 mm. on either side of the centre.

31. Prove that, whether a small-angled prism is in the position of least deviation or not, the angle of deviation is equal to $(n - 1) A$ (provided the angle of incidence is small). n = index; A = angle of prism.

32. Parallel rays of monochromatic light fall on two 10° crown glass prisms, placed base to base, the incident rays of light being parallel to the common base. On emergence a converging lens of focal length 50 cm. receives the light and forms two real images (in its focal plane) 8.68 cm. apart. Find the common index of refraction of the prisms for this kind of light. (N.B. Take $\sin A = A$.)

33. A spectrometer whose collimating lens has a focal length of 20 cm. is equipped with two slits, S_2 and S_1, 2 mm. apart, each of which is illuminated with the same monochromatic source. Each of the emergent beams is refracted through a 60° glass prism (index = 1.620 for this light) placed in the position of minimum deviation for the light coming from S_1. The emergent light is observed in a telescope, whose objective has a focal length of 30 cm. Find the distance apart of the two images observed in the focal plane of this objective. Illustrate with a good diagram.

34. In a certain Newton's Ring's arrangement, the centre is bright instead of dark. Assuming that the surfaces at the centre are separated the smallest amount necessary to make the centre bright, find the relation connecting D the diameter of the nth dark ring, the wave-length, and R the radius of curvature of the curved surface.

35. A Newton's Ring's arrangement is made by resting a convex glass surface of radius of curvature R_1 on a second convex glass surface of radius of curvature R_2. If interference fringes are observed normally with monochromatic light of wave-length, prove that

$$\lambda = \frac{D^2(R_1 + R_2)}{4nR_1R_2},$$

where D is the diameter of the nth dark ring.

36. A vertical soap (water) film formed on a rectangular framework, is illuminated with light whose wave-length in air is 5.8×10^{-5} cm. Just before the film breaks there are 10 black and 9 bright interference fringes between the top and the bottom. If the total length of the film is 12 cm. and the index air-water is 1.33 find (i) the thickness of the water film at the base just before breaking; and (ii) the angle of the wedge.

37. A Rayleigh refractometer is made with a narrow slit, a collimating lens, a double slit of width (centre to centre) 1 mm. and a telescope with objective of focal length 100 cm.

(a) Make a diagram showing the optical arrangement.

(b) When a thin flake of glass of index 1.543 is placed in the path of the monochromatic light ($\lambda = 6 \times 10^{-5}$ cm.) before one of the openings of the double slit, the fringe shift of the central fringe in the focal plane of the objective is 3.3 mm. Find the thickness of the glass flake. Show that the value of λ need not be known.

38. A Newton's Rings apparatus (in which a convex surface rests on a plane surface) is illuminated with monochromatic light of wave-length 6.0×10^{-5} cm. It is found that the difference between the squares of the diameters of any two successive bright fringes is constant and equal to 0.012 (measurement in centimetres). Find the radius of curvature of the convex surface.

39. A small distant object emitting light of wave-length 5.8×10^{-4} mm. is viewed through a telescope before the objective of which is placed a narrow rectangular aperture of width 0.58 mm. If the objective has a focal length of 100 cm., find the difference in phase between the disturbances coming from the edges of the aperture to a point P in the focal plane of the objective 0.5 mm. away from the principal focus.

40. The grating elements of two transmission grating are of equal size, but in one the width of a single opening is twice that in the other. What differences, if any, would you expect to find in the spectra?

41. When plane waves of monochromatic light fall on a small circular aperture, of diameter 1 mm., and the emergent diffracted light is observed on a screen, the screen has to be moved 12.5 cm. from the first position where the centre of the light pattern is black to the second similar position. Find the wave-length of the light used.

42. Parallel rays of two nearly equal wave-lengths λ and $\lambda + \triangle\lambda$ fall normally on a grating with element of width S. If θ is the angular distance of the first order image of λ from the

centre, prove that $\triangle\theta$, the angular separation of the first order images of the two wave-lengths, is given by

$$\triangle\theta = \frac{\triangle\lambda}{S \cos\theta}.$$

43. A source emitting two wave-lengths in the neighborhood of 6000 A illuminates the narrow slit of a spectrometer. When the light emerging from the prism on the spectrometer table is observed in a telescope with micrometer eyepiece, it is found that the centres of the images of the two spectral lines are 0.5 mm. apart for full aperture of the objective. With a rectangular narrow slit in front of the objective, the two spectral lines fuse into each other for a width of 0.6 mm. Find the focal length of the objective.

44. Plane waves of light of wave-length 5×10^{-5} cm. fall normally on *either* a small circular object of radius 0.5 mm., *or* an opaque screen with a small circular hole of this same radius. If for each case the light is observed at a point P 50 cm. beyond the object (or the hole), this point being on a line which is a normal passing through the centre of the object or the hole, prove that the amplitude of the light at P, in the case of the hole, is approximately twice that observed in the case of the object.

45. A parallel beam of plane polarized light for which $\lambda = 6 \times 10^{-5}$ cm. is incident normally on a thin quartz wedge whose angle is $\frac{1}{2}°$. The wedge is cut so that the optic axis is parallel to the thin edge of the wedge. Find the distance between any two interference fringes seen on the face of the wedge.

(Take $n_E = 1.5533$; $n_O = 1.5442$.)

46. Plane polarized light of wave-length 6×10^{-5} cm., falls normally on a piece of doubly refracting crystal of thickness 0.00025 cm., cut with optic axis parallel to the surface. If the vibration plane of the incident light makes an angle of 30° with the principal plane of the crystal and $n_E = 1.57$, $n_O = 1.55$, find (i) the phase difference between O and E beams on emergence; (ii) the ratio of the amplitudes of the E and the O beams; (iii) the nature of the resulting vibration.

47. A horizontal beam of light after passing through a Nicol prism traverses a tank containing water with scattering particles

in suspension. An observer views, through a second Nicol, the scattered light in a (horizontal) direction at right angles to the beam. Find the ratio of the intensity observed when the principal plane of each of the two Nicols is vertical, to that when the principal plane of each makes an angle of 30° with a vertical plane.

48. A thin piece of quartz in the form of a small-angled wedge is placed between crossed Nicol prisms in a parallel beam of monochromatic light ($\lambda = 6 \times 10^{-5}$ cm.), with its faces approximately normal to the light. Interference fringes are observed on the face of the quartz 7.6 mm. apart. If the optic axis is parallel to the refracting edge of the wedge, find the angle of the wedge.

(Take $n_E = 1.553$; $n_O = 1.544$.)

49. Plane polarized light falls normally on a thin piece of doubly refracting crystal, so cut that the optic axis lies in the surface of the crystal. On emergence the light is elliptically polarized in such a way that one of the axes of the ellipse is parallel to the principal plane of the crystal. What is the least thickness of the crystal for which this is the case?

Take $n_E = 1.556$; $n_O = 1.546$; $\lambda = 5 \times 10^{-5}$ cm.

50. A horizontal beam of plane polarized light is observed after passing through two Nicol prisms, *A* and *B*, in succession. If the vibration plane of the original beam makes an angle of 10° with the principal plane of the first Nicol, and if the principal planes of the two Nicols are originally parallel, find the ratio of the intensity of the emergent beam when Nicol *A* is rotated through 20° (away from the vibration plane of the original beam) with Nicol *B* stationary, to that when Nicol *A* is left stationary and *B* is rotated through 20° in the same direction.

51. The Lyman series for the hydrogen spectrum is given by

$$= 109678\ (1 - 1/m^2).$$

Find the ionization potential of the hydrogen atom.

52. Given that the Rydberg constant for hydrogen is 109678, $e = 4.77 \times 10^{-10}$ (e.s. units), $h = 6.55 \times 10^{-27}$, find, (i) the ionization potential of the hydrogen atom, (ii) the excitation potential necessary to bring out $\lambda 6563$, the first member of the *Balmer* series.

53. Two important lithium spectral series are given by

$$\bar{\nu} = 43486 - m^2P_{\frac{1}{2}}.$$
$$\bar{\nu} = 28582 - m^2S_{\frac{1}{2}}.$$

Find (i) the wave-length which you would expect to show resonance radiation (or to be readily absorbed); (ii) the corresponding radiating potential; (iii) the ionization potential of the lithium atom.

ANSWERS TO SUPPLEMENTARY PROBLEMS

1. $y = 6 \sin\left(\frac{\pi x}{100} - 72^\circ\right)$. **2.** $y = -6 \sin \frac{\pi x}{50}$.

(i) 5.66 cm. (ii) − 1.46 cm.; − 15°. **4.** 300 cm.

(i) 8.55. (ii) $y = 8.55 \sin\left(\frac{2\pi t}{T} - 20\right.$

6. (a) − 93°; + 87°. (b) 100 cm. (c) $y = -6 \sin \frac{\pi x}{50}$.

7. + 10.4. **8.** (i) + 150°; (ii) + 195°; (iii) 36 cm.; (iv) − 5.20.
9. 200 cm. per sec. **10.** (i) 2 units; (ii) 24 units; (iii) 7 units.
11. − 114 cm. from second lens. **12.** 0.67 cm.
13. 20 cm. from lens. **14.** − 429 cm. from lens (virtual).
15. 40 cm. from surface. **16.** 3 cm. **17.** 1.125 cm.
18. 4.57 cm. **19.** 1.33 cm. **20.** 34.1 cm. from lens.
21. 10 cm. **22.** 8.33. **23.** 5. **24.** 5. **25.** 130° 36′.
26. 0.03. **27.** 4 min. **28.** 1.2 cm. **29.** − 30.4 cm.
30. $\frac{\pi^2}{4}$. **32.** 1.5. **33.** 2.97 mm. **34.** $\lambda = \frac{D^2}{2R(2n-1)}$.
36. (i) 1.9×10^{-4} cm.; (ii) 0.16×10^{-4} radian.
37. 0.0061 mm. **38.** 50 cm. **39.** 180°. **41.** 0.00005 cm.
43. 50 cm. **45.** 7.6 mm. **46.** (i) 30°; (ii) $\sqrt{3} : 1$.
47. $\cos^4 30^\circ$. **48.** 30 min. **49.** 0.0125 mm.
50. $\frac{\cos^2 30^\circ}{\cos^2 10^\circ}$. **51.** 13.5 volts. **52.** (i) 13.5 volts; (ii) 12.0 volts.
53. (i) 6708 A; (ii) 1.84 volts; (iii) 5.4 volts.

INDEX